

Regional-scale phenology modeling based on meteorological records and remote sensing observations

Xi Yang,^{1,2} John F. Mustard,¹ Jianwu Tang,² and Hong Xu³

Received 3 February 2012; revised 26 July 2012; accepted 1 August 2012; published 14 September 2012.

[1] Changes of vegetation phenology in response to climate change in the temperate forests have been well documented recently and have important implications on the regional and global carbon and water cycles. Predicting the impact of changing phenology on terrestrial ecosystems requires an accurate phenology model. Although species-level phenology models have been tested using a small number of vegetation species, they are rarely examined at the regional level. In this study, we used remotely sensed phenology and meteorological data to parameterize the species-level phenology models. We used a remotely sensed vegetation index (Two-band Enhanced Vegetation Index, EVI2) derived from the Moderate Resolution Spectroradiometer (MODIS) 8-day reflectance product from 2000 to 2010 of New England, United States to calculate remotely sensed vegetation phenology (start/end of season, or SOS/EOS). The SOS/EOS and the daily mean air temperature data from weather stations were used to parameterize three budburst models and one senescence model. We compared the relative strengths of the models to predict vegetation phenology and selected the best model to reconstruct the “landscape phenology” in New England from year 1960 to 2010. Of the three budburst models tested, the spring warming model showed the best performance with an averaged Root Mean Square Deviation (RMSD) of 4.59 days. The Akaike Information Criterion supported the spring warming model in all the weather stations. For senescence modeling, the Delpierre model was better than a null model (the averaged phenology of each weather station, averaged model efficiency = 0.33) and has a RMSD of 8.05 days. A retrospective analysis using the spring warming model suggests a statistically significant advance of SOS in New England from 1960 to 2010 averaged as 0.143 days per year ($p = 0.015$). EOS calculated using the Delpierre model and growing season length showed no statistically significant advance or delay between 1960 and 2010 in this region. These results suggest the applicability of species-level phenology models at the regional level (and potentially terrestrial biosphere models) and the feasibility of using these models in reconstructing and predicting vegetation phenology.

Citation: Yang, X., J. F. Mustard, J. Tang, and H. Xu (2012), Regional-scale phenology modeling based on meteorological records and remote sensing observations, *J. Geophys. Res.*, 117, G03029, doi:10.1029/2012JG001977.

1. Introduction

[2] Long term phenological observations from the northern hemisphere provided evidence that climate change is driving shifts in vegetation phenology [Fitter and Fitter, 2002; Schwartz *et al.*, 2006]. Vegetation start-of-season (SOS) and end-of-season (EOS) are two key phenological phases (i.e.,

phenophases) to determine the plant growing season length (GSL), which is an important parameter in terrestrial carbon cycle in the temperate deciduous forest [Churkina *et al.*, 2005; Dragoni *et al.*, 2011; Piao *et al.*, 2007; Picard *et al.*, 2005; Richardson *et al.*, 2010]. Changes in phenology also feedback on the climate system through the nutrient cycle, the water cycle, the surface energy budget and the production of biogenic volatile organic compounds (BVOCs) [Peñuelas *et al.*, 2009; Schwartz, 1996]. At the community level, shifts in the phenology of related species (e.g., flowering plants and pollinators) might cause mismatches in reproductive timing and failure to produce offspring [Bradshaw and Holzapfel, 2006]. Better modeling of vegetation phenology is thus critical to predict how the ecosystem will respond to the future climate.

[3] Both SOS and EOS are controlled by various environmental factors. It is widely accepted that leaf budburst in temperate forests is mainly driven by temperature [Cannell and Smith, 1983; Hänninen, 1990; Peñuelas and Fillela,

¹Department of Geological Sciences, Brown University, Providence, Rhode Island, USA.

²The Ecosystem Center, Marine Biological Laboratory, Woods Hole, Massachusetts, USA.

³Department of Soil, Water and Climate, University of Minnesota, St. Paul, Minnesota, USA.

Corresponding author: J. F. Mustard, Department of Geological Sciences, Brown University, Providence, RI 02912, USA. (john_mustard@brown.edu)

©2012. American Geophysical Union. All Rights Reserved. 0148-0227/12/2012JG001977

2001]. However, the time interval in which air temperature is effective is still widely debated: spring (here refers to Jan–Jun) only (e.g., [Chuine et al., 1998; Richardson et al., 2006]) or winter and spring (e.g., [Chuine et al., 1999; Kramer, 1994; Vitasse et al., 2011]). Other factors such as photoperiod [Partanen et al., 1998], precipitation and nitrogen deposition [Cleland et al., 2007] are considered to have less effects on the budburst at the ecosystem level: a meta-analysis, which summarizes the environmental cues on spring phenology, suggested that of 325 species surveyed, 278 species are cued by temperature while only 2 species are cued by photoperiod (35 species by precipitation, which is one of the main drivers in tropical forests [Reich, 1995]) [Pau et al., 2011]. On the contrary, there is still no consensus on the driving factors of fall senescence (defined as the leaf coloring), possibly due to the lack of understanding of the senescence process and limited availability of senescence data [Delpierre et al., 2009]. The possible factors that might control fall senescence include summer temperature [Estrella and Menzel, 2006] and photoperiod [Keskitalo et al., 2005].

[4] Based on these analyses, phenology models have been developed from species to global levels. At the species level, phenology models for both budburst and senescence were developed based on controlled experiments and have been tested using phenological observations of dominant tree species in Europe [Chuine et al., 1998; Häkkinen et al., 1998; Hänninen, 1990] and North America [Chuine et al., 2000; Richardson et al., 2006]. It should be noted that the budburst models could not only simulate budburst, but also other stages in the spring canopy development. Budburst models assume a linear relationship between the rate of growth (e.g., the rate of increase in mean leaf area) and temperature above a given threshold (“growing degree days” (GDD)). When a certain temperature accumulation threshold (“critical forcing temperature” (F^*)) is reached, budburst occurs. Some models such as the spring warming model [Hunter and Lechowicz, 1992] assume that only spring and summer temperature (Jan–Jun) has an impact on the budburst while the other models such as sequential [Sarvas, 1974] and parallel models [Kramer, 1994; Landsberg, 1974] require a cold winter – the number of days with daily temperature below a certain threshold (e.g., 2°C) reaches the chilling requirement (e.g., 15 days) – to initiate the spring temperature accumulation process while the spring warming model implicitly assumes that this winter “chilling requirement” is always fulfilled. There are complex models such as Promoter-Inhibitor Model (PIM) [Schaber and Badeck, 2003] and UniChill model [Chuine, 2000], which require at least seven years of data to avoid model over-fitting. Senescence is defined as the process of leaf coloring (red or yellow). Estrella and Menzel [2006] used more than 50 years of in-situ observations of autumn senescence of four common tree species in Germany to test the relationship between commonly used criteria (e.g., summer temperature, solar radiation) and leaf senescence, and found that the criteria were not sufficient to explain the variation in leaf senescence. In contrast, Delpierre et al. [2009] found that a combination of temperature and photoperiod was sufficient to predict the senescence date of Sessile Oak (*Quercus petraea* (Matt.) Liebl.) and European beech (*Fagus sylvatica*). Vitasse et al. [2011] extended this method to four dominant species in European temperate forests and found that the senescence

model has good predictability for three species: *Quercus petraea*, *Acer pseudoplatanus* and *Fagus sylvatica*.

[5] At the regional and global level, phenology models are mainly used as submodels in the terrestrial biosphere models; Most of these phenology models are empirical, using either prescribed date, or a single temperature threshold or GDD without parameter optimizations using phenological observations (for details of the models see Richardson et al. [2012]). Thus, due to their inaccurate characterization of vegetation dynamics, these phenology models applied to the regional or global scale might underestimate or overestimate the effects on the biosphere [Randerson et al., 2009]. Comparison of phenology models in 14 terrestrial biosphere models suggests that none of the models succeeded in capturing the phenology in terms of leaf area index (LAI) or carbon fluxes, and most of the models predicted an earlier SOS and later EOS, resulting in overestimation of gross ecosystem photosynthesis by $\sim 20\%$ [Richardson et al., 2012]. Comparing to these empirical models, species-level models such as the spring warming or parallel models are better supported by phenological observations [Migliavacca et al., 2012]. The species-level models are rarely tested at the regional level [Fisher et al., 2007; Picard et al., 2005]. Remote sensing provides a way to monitor several key phenological phases including leaf expansion and leaf coloring at the landscape scale (e.g. [Fisher and Mustard, 2007; Fisher et al., 2006; Morisette et al., 2009; Zhang et al., 2003]). Fisher et al. [2007] used 5 years of remotely sensed phenology and climate data to parameterize the spring warming model in New England, USA. The work was limited by the short time span of the remote sensing data and thus the model was not well-fit at each individual weather station (5 years vs. 3–5 parameters per model). Now we have 11 years (as of 2010) of remotely sensed data from MODIS that allows for a more robust model fitting with data at individual stations. We assume that each weather station records the climate data for a unique mixture of vegetation, thus phenology model parameters are spatially different, but for the same station, the model parameters are temporally invariant. In the present study, we choose New England in northeastern United States as our study area to address the following questions: (1) Can species-level budburst and senescence models predict the remotely sensed phenology better than a null model (i.e., the eleven-year-averaged remotely sensed phenology of a given location)? Which budburst model is the best? (2) Is there a trend in phenology in New England in the past 50 years?

2. Methods

2.1. Study Area

[6] The study area (40°N – 44°N , 69°W – 76°W) encompasses southern New England extending west to east from central New York to Martha’s Vineyard, Massachusetts and north to Vermont and Maine (Figure 1). It is considered as the “tension zone” between two distinct hardwood forest communities: in the north are mainly beech, birches, and maples; in the south are mainly oaks, chestnuts and hickories. The major tree species are white pine (*Pinus strobus*), yellow birch (*Betula alleghaniensis*), red maple (*Acer rubrum*), red oak (*Quercus rubra*), and white oak (*Quercus alba*) [Cogbill et al., 2002].

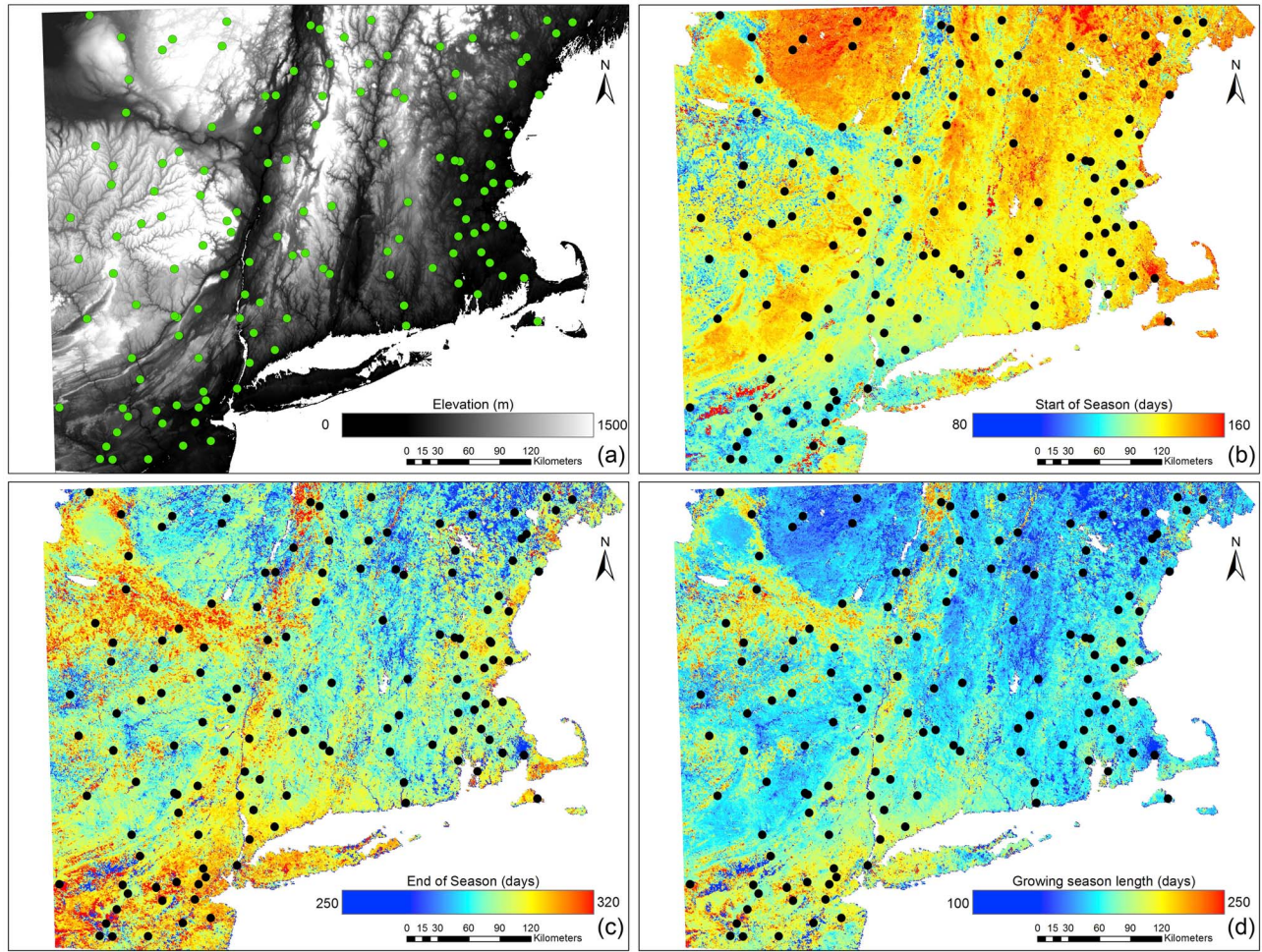


Figure 1. Distribution of weather stations in the study area and with the following properties as the background: (a) the elevation; (b) start of season (SOS); (c) end of season (EOS); and (d) growing season length (GSL). SOS, EOS and GSL were calculated using the MODIS Two band Enhanced Vegetation Index (EVI2) time series in year 2002 as an example. For calculation methods, see Figure 2. The green and black dots indicate locations of the weather stations.

2.2. Remotely Sensed Phenology and Spatial Weighting

[7] We used the remotely sensed data to estimate the vegetation phenology of the study area. The 8-day 500 m MODIS surface reflectance products (code: MOD09A1) of the study area were acquired from the NASA LPDAAC (<https://lpdaac.usgs.gov/>). The two-band Enhanced Vegetation Index (EVI2) [Jiang *et al.*, 2008] was calculated for each pixel from 2000 to 2010 as follows (equation (1), ρ_{NIR} is the near-infrared band reflectance, ρ_{RED} is the red band reflectance):

$$EVI2 = 2.5 \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + 2.4\rho_{RED} + 1}, \quad (1)$$

Unlike EVI, EVI2 uses only the near-infrared and red bands from MODIS, making it possible to extend the use of EVI2 to sensors like AVHRR. When atmospheric effects are minimal, the difference between EVI and EVI2 is insignificant when tested over various land cover/use types and different times of the season, and EVI and EVI2 do not

become saturated even when LAI exceeds 5 [Jiang *et al.*, 2008]. Band quality files and state flags with the data were used to screen the cloudy days, water surface and other erroneous pixels. Then the EVI2 time series were processed using the Savitzky-Golay Filter, which has been used to smooth vegetation index time series with erroneous spikes due to clouds [Chen *et al.*, 2004]. The smoothed time series were fitted using a double-logistic function (equation (2)) [Fisher and Mustard, 2007; Fisher *et al.*, 2006]:

$$v(t) = v_{\min} + v_{\text{amp}} \left(\frac{1}{1 + e^{m_1 - n_1 t}} - \frac{1}{1 + e^{m_2 - n_2 t}} \right), \quad (2)$$

where $v(t)$ is the EVI2 at time t , v_{\min} and v_{amp} are the minimum and amplitude values of a single year and the parameters m_1 , n_1 , m_2 , and n_2 control the shape of the curve (Figure 2). The curve-fitting procedure used was MPFIT, a robust non-linear least-square fitting method [Markwardt, 2008, downloadable from <http://purl.com/net/mpfit>]. Specifically, $t = m_1/n_1$ is the SOS and $t = m_2/n_2$ is the EOS. The

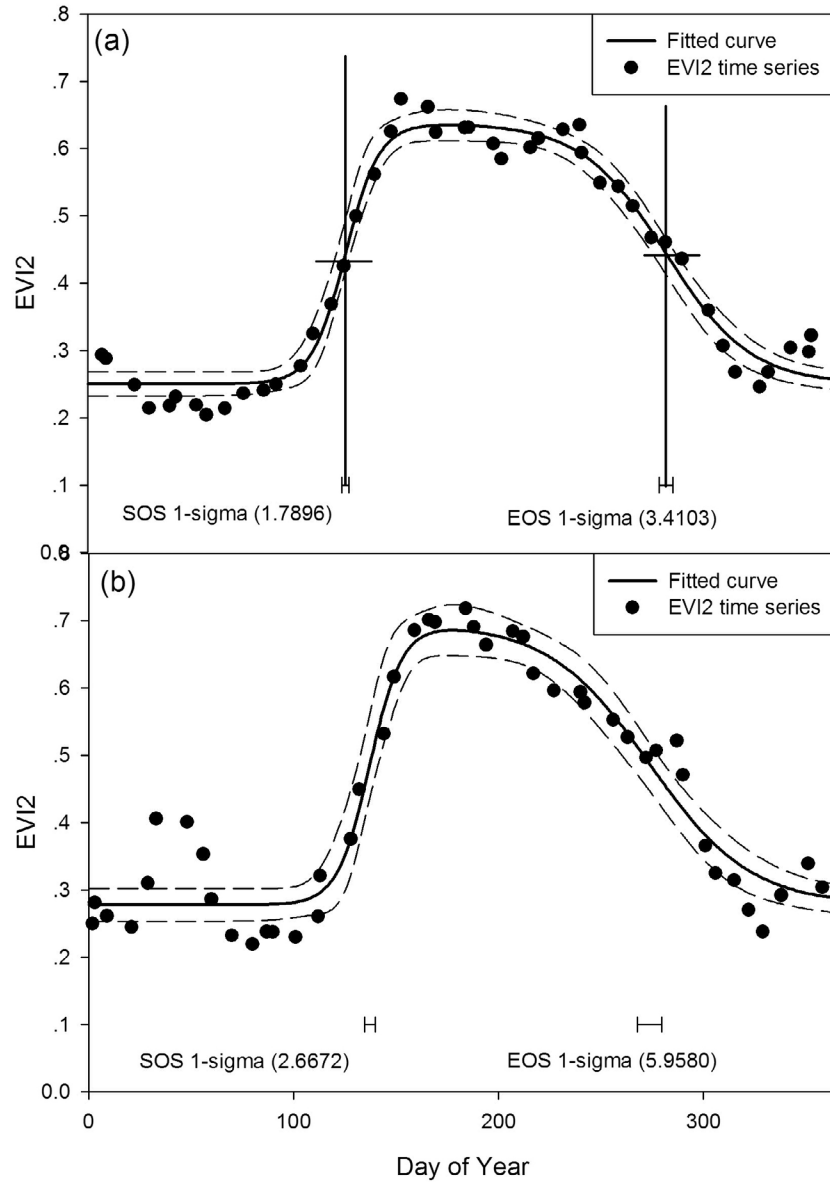


Figure 2. Examples of curve-fitting using the double-logistic function. The solid curves are the fitted double-logistic function, with the 95% confidence interval on both sides of the curves (dash lines). The black dots are the EVI2 time series. Two vertical lines in panel (a) show the dates that were calculated as SOS and EOS. The whiskers below the fitted curves are uncertainties in SOS and EOS. The unit of the uncertainty is day.

calculated SOS and EOS are the days that the vegetation index increases/decreases to the halfway point between the maximum and minimum value. This method is considered to be less sensitive to the understory species green-up [Fisher *et al.*, 2006], which is often earlier than that of the overstory dominant species [Richardson and O'Keefe, 2009].

[8] We assessed the uncertainty of remotely sensed phenology modeling in two ways: ground validation (see section 2.3) and uncertainty evaluation of parameters related to SOS and EOS. We used the 1-sigma value of m_1 , n_1 ($\sigma_{m_1}, \sigma_{n_1}$) and 1-sigma value of m_2 , n_2 ($\sigma_{m_2}, \sigma_{n_2}$) to assess the uncertainty in SOS and EOS, respectively. Since m_1 and n_1 (m_2 and n_2) are highly correlated (ρ_{m_1, n_1} and ρ_{m_2, n_2} , correlation between m_1 , n_1 ; and m_2 , n_2 , data not shown), the propagation

of the uncertainty to SOS and EOS should be described as σ_{SOS} and σ_{EOS} [Taylor, 1997]:

$$\sigma_{SOS} = SOS \cdot \sqrt{\left(\frac{\sigma_{m_1}}{m_1}\right)^2 + \left(\frac{\sigma_{n_1}}{n_1}\right)^2 - 2 \frac{\sigma_{m_1} \sigma_{n_1}}{m_1 n_1} \rho_{m_1 n_1}} \quad (3)$$

$$\sigma_{EOS} = EOS \cdot \sqrt{\left(\frac{\sigma_{m_2}}{m_2}\right)^2 + \left(\frac{\sigma_{n_2}}{n_2}\right)^2 - 2 \frac{\sigma_{m_2} \sigma_{n_2}}{m_2 n_2} \rho_{m_2 n_2}}, \quad (4)$$

where SOS and EOS are the estimated values of a year from equation (2); (m_1, m_2, n_1, n_2) are the estimated best values of the parameters.

[9] The weather stations record the temperature of their surrounding area. However, the satellite pixels around the weather stations should not be considered equally valid due to different land use/cover and elevation [Fisher *et al.*, 2007]. Thus, we gave weights to the 7×7 grids surrounding the weather station. The weights are based on the vegetation cover of each pixel, the distance with the central pixel (i.e., the location of weather station), and the elevation difference between the pixel and the central pixel, and the water mask. The averaged SOS and EOS were then calculated as the weighted average of the 49 pixels in the grid.

[10] A pure deciduous forest in New England has an annual maximum EVI2 value (v_{\max}) close to ~ 0.8 , and an annual minimum EVI2 value (v_{\min}) close to ~ 0.0 . Thus the Cartesian distance in the $v_{\max} - v_{\min}$ space indicates the ‘deciduousness’ of a pixel: the smaller the value is (thus higher W_{DC}), the closer the pixel is to be considered as a deciduous forest [Fisher *et al.*, 2006]. If v_{\max} is greater than 0.8, then v_{\max} was set to be equal to 0.8. The deciduousness (W_{DC}) of a pixel is (non-deciduous pixels are thus very low in W_{DC} and readily dismissed):

$$W_{DC} = 1 - \sqrt{(v_{\min})^2 + (v_{\max} - 0.8)^2}. \quad (5)$$

The horizontal distance weight (W_{HD}) decreases from 1 to 0 with a lapse rate of 1/7. Calculated as the Cartesian distance between the central pixel ($i = 3$ and $j = 3$, where i and j are the horizontal and vertical coordinates),

$$W_{HD} = 1 - (1.0/7) \times \sqrt{(i - 3)^2 + (j - 3)^2}. \quad (6)$$

The vertical distance weight (W_{VD}) was calculated as the difference of elevation between the central point and any point on the grid with a lapse rate of 0.02 units m^{-1} :

$$W_{VD} = 1 - 0.02 \times |\text{altitude}[i, j] - \text{altitude}[3, 3]|. \quad (7)$$

The water mask weight (W_{water}) equals to 0 when the pixel is recognized as water by MODIS state flag, otherwise W_{water} equals to 1. The total weight (W) of each pixel on the grid is:

$$W = W_{DC} \times W_{HD} \times W_{VD} \times W_{water}. \quad (8)$$

2.3. Ground Validation

[11] We used the phenology records in Harvard Forest (42.53°N–42.54°N, 72.18°W–72.19°W) to validate remotely sensed SOS and EOS. Harvard Forest is a mixed forest dominated by red maple (*Acer rubrum*) and red oak (*Quercus rubra*). Spring canopy developments of 33 species were recorded since 1990 at an interval of 3–7 days (after 2002, the total number of species in spring was reduced to nine); fall canopy developments were observed since 1991 (except 1992) and the number of species is 14 since 2002 [O’Keefe, 2000]. Eleven years of observations of *A. rubrum* (5 individuals were observed) and *Q. rubra* (4 individuals were observed) were used to compare with remotely sensed SOS and EOS. Each time, phenological metrics were recorded as the percentage comparing to the total leaves on that tree (three spring metrics and two fall metrics, 0% to

100%): BBRK (percentage of broken buds), L75 (percentage of leaves at 75% of their total size), L95 (percentage of leaves at 95% of their total size), LCOLOR (percentage of leaves that have changed color, notice that the leaves are those remaining on the tree) and LFALL (percentages of leaves that have fallen). All of the metrics for each individual were fitted using a sigmoid curve [Fisher *et al.*, 2007, equation (9)]:

$$PM(t) = PM_{\min} + PM_{\text{amp}} \left(\frac{1}{1 + e^{m_1 - n_1 t}} \right), \quad (9)$$

where $PM(t)$ is the phenological metrics at time t , PM_{\min} and PM_{amp} is the minimum and amplitude values of the above metrics of a single year. The parameters m_1 , n_1 control the shape of the curve. Similar to section 2.3, we calculated the date ($t = m_1/n_1$) when the metrics reach halfway between minimum and maximum to compare with the remotely sensed phenological metrics of the Harvard Forest pixel from 2000 to 2010. The date of those metrics should be interpreted as “the date when 50% of the leaves on the tree reach that stage (for example, budburst or reaches 75% of the full leaf size).”

2.4. Climate Data

[12] Daily temperature and photoperiod are used as climate drivers of the phenology models. Daily maximum and minimum temperature from 1999 to 2010 were acquired from NOAA National Climate Data Center (www.ncdc.noaa.gov/oa/ncdc.html). These data were processed in the following steps: first, the daily mean temperature was calculated as the average of the daily maximum and minimum temperature [Fisher *et al.*, 2007]; second, stations with more than 15% of the data flagged as missing (“–99999” in the original dataset) were discarded. The remaining missing data were replaced by the interpolation of nearby stations using the Linear Lapse Rate Adjustment (LLRA) [Dodson and Marks, 1997]; third, we compiled the data from September to the next June as the dataset input for spring phenology models; data from June to December were compiled for fall phenology model. Stations with 5 or more years of data were included in the dataset. Stations located in the airports and croplands were manually excluded based on the examination of Google Earth images from 2000 to 2010. The total number of included weather stations is 137. In addition, we calculated the daily photoperiod for each station as a function of the latitude of the station and the day of year [Monteith and Unsworth, 2008].

2.5. Parameterization of Climatic Phenology Models

[13] The models we selected in this paper must be simple in terms of the number of parameters, since we only have 11 years of satellite data. Models such as the promotor-inhibitor model (PIM) [Schaber and Badeck, 2003] with more than 5 parameters were not considered. For budburst, both 1-phase (which only consider the spring temperature) and 2-phase models (which consider both the fall and next spring temperature) were considered in this study (Table 1). For the 1-phase model, we selected the spring warming model (SW), which accumulates growing degree days (GDDs) after a given DOY (which could vary spatially). For

Table 1. Summary of Model Equations and Notations Used in this Study, and the Temporal Range of Temperature and Photoperiod Records Required by the Models

| Model | Parameters ^a | Spring Temperature ^b | Autumn Temperature ^b | Autumn Photoperiod ^b | Equation |
|----------------|---|---------------------------------|---------------------------------|---------------------------------|--|
| Spring warming | $T_{\text{heat}}, t_0, F^*$ | ✓ | | | SW: $S_f = \sum_{t_0}^t R_f(x_t)$ where $R_f = \max(0, x_t - T_{\text{heat}})$ when $S_f \geq F^*$ budburst occurs |
| Sequential | $C_{\text{total}}, T_{\text{heat}}, T_{\text{chill}}, F^*$ | ✓ | ✓ | | SEQ: $S_c = \sum_{t_0}^t R_c(x_t)$ where $R_c = \text{binary}(\max(0, T_{\text{chill}} - x_t))$ when $S_c \geq C_{\text{total}}$ heat accumulation starts then $S_f = \sum_{t_h}^t R_f(x_t)$ where $R_f = \max(0, x_t - T_{\text{heat}})$ when $S_f \geq F^*$ budburst occurs |
| Parallel | $C_{\text{total}}, T_{\text{heat}}, T_{\text{chill}}, F^*$ | ✓ | ✓ | | PAR: $S_c = \sum_{t_0}^t R_c(x_t)$ where $R_c = \text{binary}(\max(0, T_{\text{chill}} - x_t))$ $K_m = \min(S_c/C_{\text{total}}, 1)$ and $S_f = \sum_{t_0}^t K_m \cdot R_f(x_t)$ where $R_f = \max(0, x_t - T_{\text{heat}})$ when $S_f \geq F^*$ budburst occurs |
| Delpierre | $P_{\text{start}}, T_{\text{chill}}, x, y, Y_{\text{crit}}$ | | ✓ | ✓ | DM: If $P(d) \leq P_{\text{start}}$ and $x_t \leq T_{\text{chill}}$, then $S_{\text{sen}} = \sum_{t_1}^t R_{\text{sen}}(x_t)$ where $R_{\text{sen}}(x_t) = [T_{\text{chill}} - x_t]^y \times [P(t)/P_{\text{start}}]^y$ when $S_{\text{sen}} \geq Y_{\text{crit}}$ senescence occurs ^c |

^a S_f is the accumulated heat forcing units (unit: °C); R_f is the rate of heat forcing (unit: °C/day); S_c the accumulated chilling units (unit: °C); R_c is the rate of chilling (unit: °C/day); x_t is the temperature at time t ; T_{heat} , base temperature (unit: °C) required by heat accumulation process; T_{chill} is base temperature (unit: °C) required by chilling accumulation process; t_0 is the starting date (day of year, unit: day) of accumulation; t_b is the date of budburst (day of year, unit: day); t_h is the date when the heating accumulation is completed (day of year, unit: day); t_s is the date of senescence (day of year, unit: day); F^* is the critical threshold of heating process (budburst) (unit: °C); C_{total} is critical threshold of chilling process (end of chilling, quiescence) (unit: day); S_{sen} is the accumulated forcing units for senescence (unit: °C·hour·hour⁻¹), R_{sen} is the rate of forcing (unit: °C·hour·hour⁻¹/day); Y_{crit} is the critical threshold of senescence process (senescence) (unit: °C·hour·hour⁻¹); P_{start} is the photoperiod threshold for senescence process (unit: hour); $P(t)$ is the photoperiod for day t ; x, y are parameters for the DM model. Functions: $\max()$ returns the larger value of the two in the parenthesis, $\min()$ returns the smaller value in the parenthesis, while $\text{binary}()$ return 0 if the value in the parenthesis is 0, otherwise returns 1.

^bSpring is from 1 January to 30 June. Autumn is from 1 August to 31 December.

^c $P(t)/P_{\text{start}}$ can also be written as $1 - P(t)/P_{\text{start}}$ [Delpierre et al., 2009].

the 2-phase model, the sequential model (SEQ) calculates the GDD after the chilling requirement is fulfilled by having a certain number of low temperature days (chilling days (CD)) [Chuine et al., 1998; Kramer, 1994; Landsberg, 1974; Sarvas, 1974]. The parallel model (PAR) calculates the GDD concurrently with CD accumulation, and budburst happens when both heating and chilling requirements are fulfilled [Chuine et al., 1998; Landsberg, 1974].

[14] For senescence models, we tested the Delpierre model (DM) that assumes both temperature and photoperiod control the senescence process [Delpierre et al., 2009; Vitasse et al., 2009]. The accumulation of cold degree days (CDDs) is initiated when daily temperature and photoperiod are both below their threshold. We made a change to the DM model such that both the parameter for temperature (x) and photoperiod (y) could accept any value between 0 and 2 (Table 2) instead of only 0, 1 and 2 as in Delpierre et al. [2009].

[15] The remotely sensed phenology (SOS and EOS) and the daily environmental data (temperature and photoperiod) were used to calculate the phenology model parameters (for parameters see Table 1). For each weather station, at least 5 years of data were used in the model calibration. Since we assumed that the phenology model parameters at each station is temporally invariant but spatially different from stations at other locations due to biotic factors (genotypes, species composition), models were fit individually at each weather station.

[16] Since all the models have at least three parameters, it is not computationally realistic to fully explore the parameter space [Picard et al., 2005]. We utilized a simple genetic algorithm to optimize the cost function (equation (10)). Model parameters were constrained to a range of values based on previous modeling works (e.g., [Chuine et al., 1998, 1999; Kramer, 1994]) (Table 2). All the codes were written in Interactive Data Language (for genetic algorithm source code, refer to www.ncnr.nist.gov/staff/dimeo/idl_programs.html).

$$\text{RMSD} = \frac{1}{N} \sqrt{\sum \left(t_{\text{pheno}}^{\text{observation}}(\text{year}_i) - t_{\text{pheno}}^{\text{model}}(\text{year}_i) \right)^2}, \quad (10)$$

where N is the number of years of each station, $t_{\text{pheno}}^{\text{observation}}(\text{year}_i)$ is the estimated date of a phenophase (i.e., SOS and EOS) at a given year i , $t_{\text{pheno}}^{\text{model}}(\text{year}_i)$ is the modeled date of a phenophase at a given year i .

2.6. Model Evaluation

[17] Model accuracy and efficiency were analyzed using the RMSD (equation (10)), Nash-Sutcliffe model efficiency coefficient (ME) [Nash and Sutcliffe, 1970] (equation (11)), and Akaike Information Criterion (AIC) [Akaike, 1974]. RMSD describes the difference between the modeled phenophase date and the observed phenophase date. ME compares the phenology models with the null model (i.e., only

Table 2. Ranges for the Parameters in Each Model^a

| Model | Parameter 1 | Parameter 2 | Parameter 3 | Parameter 4 | Parameter 5 |
|----------------|-----------------------------|------------------------------|--------------------|----------------------------------|------------------------------|
| Spring warming | $T_{\text{heat}}: [0, 10]$ | $t_0: [1, 100]$ | $F^*: [100, 1000]$ | | |
| Sequential | $T_{\text{heat}}: [0, 10]$ | $C_{\text{total}}: [0, 150]$ | $F^*: [100, 1000]$ | $T_{\text{chill}}: [-10, 10]$ | |
| Parallel | $T_{\text{heat}}: [0, 10]$ | $C_{\text{total}}: [0, 150]$ | $F^*: [100, 1000]$ | $T_{\text{chill}}: [-10, 10]$ | |
| Delpierre | $T_{\text{chill}}: [5, 30]$ | $x: [0, 2]$ | $y: [0, 2]$ | $Y_{\text{crit}}: [1000, 15000]$ | $P_{\text{start}}: [10, 16]$ |

^aRefer to Table 1 for the parameter acronyms.

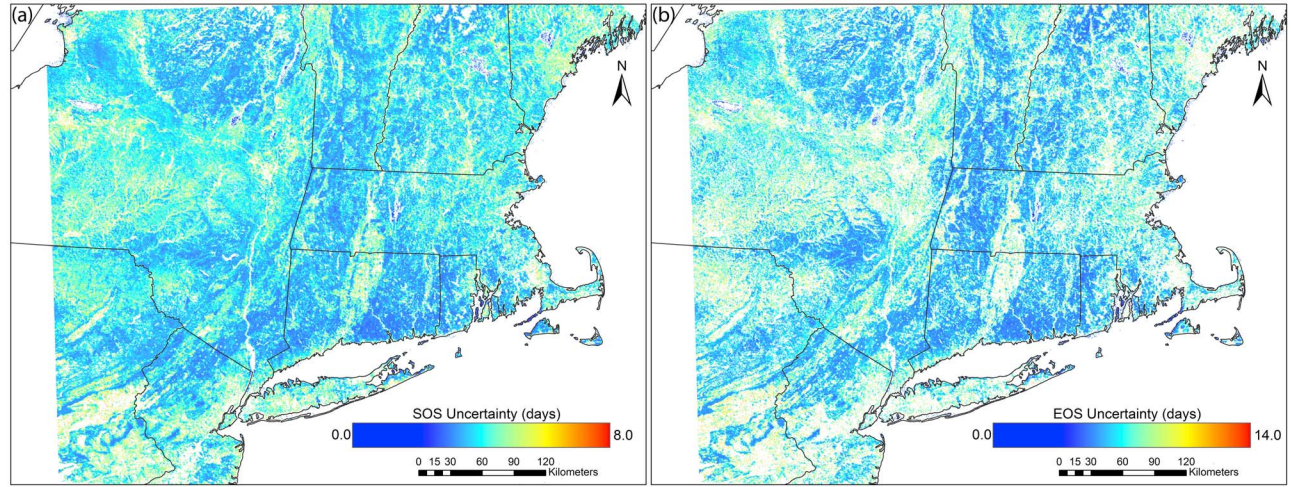


Figure 3. The cross-year averaged (2000–2010) uncertainties of SOS and EOS in the study area. (a) SOS; (b) EOS. The black solid lines are the state boundaries.

calculating the interannual variation). ME was calculated with the following equation:

$$ME = 1 - \frac{\sum (t_{pheno}^{observation}(year_i) - t_{pheno}^{model}(year_i))^2}{\sum (t_{pheno}^{observation}(year_i) - \overline{t_{pheno}^{observation}(year_i)}})^2}, \quad (11)$$

where $\overline{t_{pheno}^{observation}(year_i)}$ is the averaged date of budburst or senescence. A positive ME means the models are better than a null model.

[18] While the best model should have the as low an RMSE as possible, it is equally important that the data are fit with the fewest model parameters (“Occam’s razor”) [Burnham and Anderson, 2002]. AIC takes into account the goodness-of-fit as well as the complexity of the model. When the number of parameters (p) is large comparing to the sample size (n) (generally when $n/p < 40$), the small sample AIC should be used (AICc) [e.g., Migliavacca et al., 2012]:

$$AIC_c = n \log \sigma^2 + 2p + \frac{2p(p+1)}{n-p-1}, \quad (12)$$

where n is the number of observations, σ is the RMSD, p is the number of parameters. The model with the lowest AICc is considered to be the best model among the candidates. The difference of AICc scores between the best model and the other models, ΔAIC_c , is a measure of relative strength of the models compared to the best model. If $\Delta AIC_c < 2$, then the model is considered to be close to the best model, while if $\Delta AIC_c > 6$, then the model is 20 times less likely to be the best model [Migliavacca et al., 2012].

2.7. Retrospective Analysis

[19] Based on these metrics, the best budburst model and senescence model for the study area was identified and we chose the stations with ME higher than 0.4 for a retrospective analysis. Climate data from year 1960 to 2010 were the input to the calibrated models to derive the SOS, EOS and subtract EOS with SOS to get GSL in each year. The SOS, EOS and GSL were then averaged across the region and

linear regressions of these phenophases against year were calculated.

3. Results

3.1. Remotely Sensed Phenology and Uncertainty Analysis

[20] Figure 1 shows the spatial distribution of SOS, EOS and GSL in 2002, which is similar to the other analysis for this region [Fisher and Mustard, 2007; Zhang et al., 2003]. Spatial variations in SOS and EOS show a coastal-continental gradient with altitude as a controlling factor. The late SOS in the upper Cape Cod, Martha’s Vineyard and Long Island are mainly due to the concentration of scrub oak (*Quercus ilicifoli*) in these areas (data not shown) [Foster et al., 2002]. Scrub oak also showed an earlier EOS in the same areas (Figure 1c), together resulting in a shorter GSL.

[21] We assessed the quality of remotely sensed phenology by (1) evaluating the uncertainty of curve-fitting; (2) comparing with ground observations in Harvard Forest. Figure 2 shows two examples of curve-fitting and the uncertainties in the estimates of SOS and EOS at the MODIS pixel covering the Harvard Forest (42.535N, 72.185W). Figure 2a is year 2010 with a low scatter and Figure 2b is year 2007 with a higher scatter. The uncertainty in SOS is smaller than that of EOS. Figure 3 shows the spatial distribution of the averaged (2000–2011) uncertainties of SOS and EOS, and both shared a similar spatial pattern: the uncertainties are generally lower along the south shore, in Adirondack Mountains and in Vermont and West Massachusetts. The uncertainties are higher over croplands on the southwest corner. The averaged uncertainty of SOS (2000–2010) of the whole study area in 2.571 days with standard deviation of 0.808 days; The averaged uncertainty of EOS (2000–2010) of the whole study area is 4.458 days with standard deviation of 1.598 days. Figure 4 shows the comparison between remotely sensed phenology and Harvard Forest ground observations. Those metrics should be interpreted as when 50% of the leaves on the tree reach the state, for example, 50% of the leaves on the tree reach their 75%

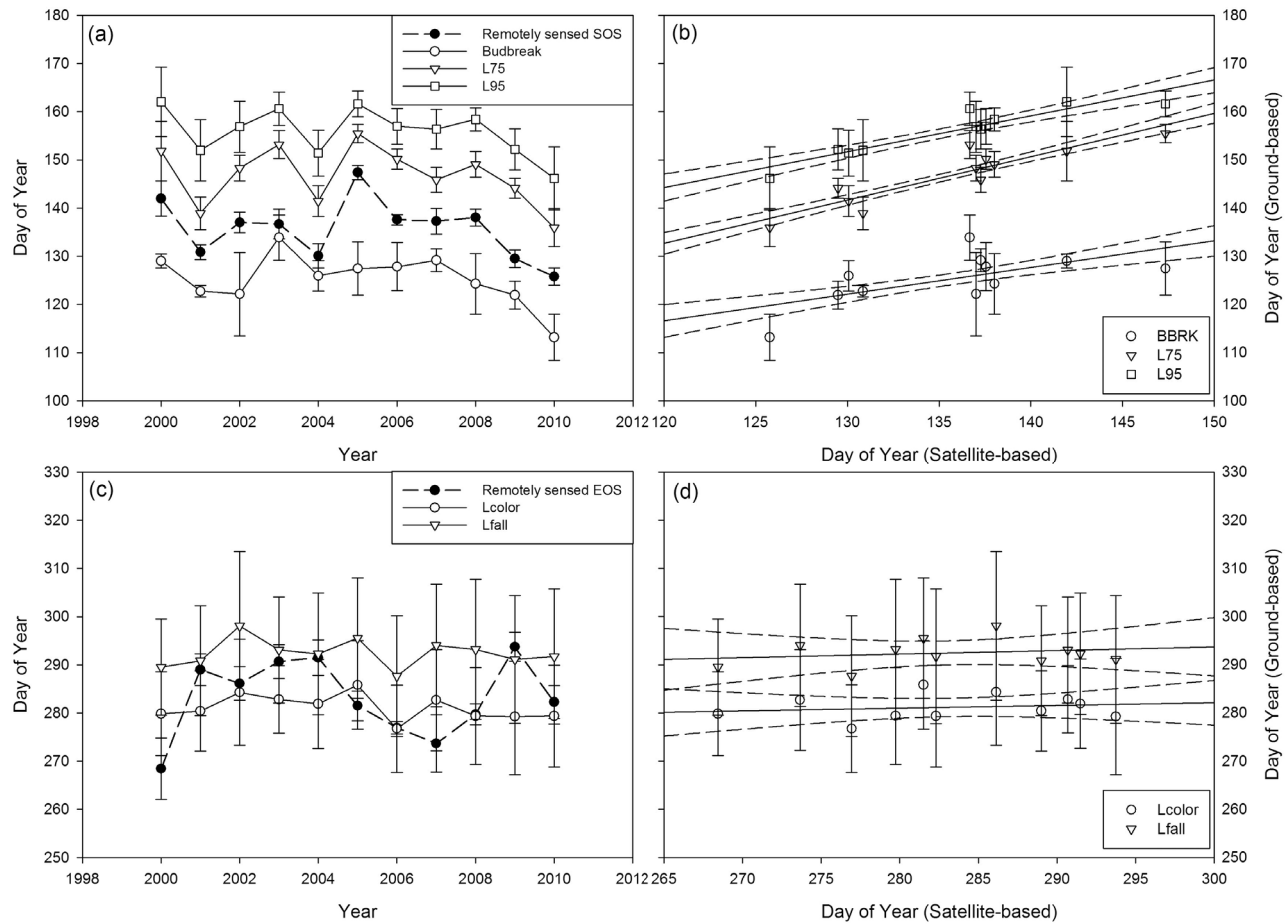


Figure 4. Comparison between remotely sensed phenology and Harvard Forest ground observations. (a) SOS and the date when spring phenological metrics (BBRK, L75 and L95) reaches 50% (see text for explanation); (b) Linear regressions between SOS and spring phenological metrics: $L95 = 0.7431 \times SOS + 55.0645$ ($r^2 = 0.4424$); $L75 = 0.9010 \times SOS + 24.5235$ ($r^2 = 0.6428$); $BBRK = 0.5539 \times SOS + 50.0835$ ($r^2 = 0.2314$). The dashed lines are 95% confidence intervals; (c) EOS and fall phenological metrics (LCOLOR and LFALL); and (d) linear regressions between EOS and fall phenological metrics (not significant). The whiskers in (a) and (c) for SOS and EOS are the uncertainties as calculated in Figure 2. The whiskers of the spring and fall phenological metrics are the standard deviation calculated from the 9 individuals (5 Red Maple and 4 Red Oak). The dashed lines are 95% confidence intervals.

size comparing to the full size (L75). L75 ($r^2 = 0.6428$) is best in tracking the remotely sensed SOS comparing to L95 ($r^2 = 0.4424$) and BBRK ($r^2 = 0.2134$). Due to the large variations of fall phenological metrics, the correlation between remotely sensed EOS and LFALL/LCOLOR is not statistically significant.

3.2. Model Performance

[22] Among the three budburst models we tested, the spring warming model showed the best performance in terms of RMSD, AIC_c and ME. The budburst models showed an average RMSD less than 5 days (Table 3). The averaged RMSD and R^2 values of the three models are close

Table 3. Summary of the Root Mean Square Deviation (RMSD), Model Efficiency (ME) and the Small Sample-Corrected Akaike Information Criterion (AIC_c) of the Models^a

| Category | Model | RMSD | ME | Median AIC_c | Best Model | $\Delta AIC_c < 2$ | $\Delta AIC_c < 6$ |
|---------------|----------------|--------------------|--------------------|----------------|------------|--------------------|--------------------|
| Spring models | Spring warming | 4.59 (2.14, 7.78) | 0.49 (0.02, 0.84) | 0.00 | 128 | 128 | 128 |
| | Sequential | 4.91 (2.24, 8.14) | 0.39 (−0.28, 0.84) | 8.45 | 0 | 0 | 28 |
| | Parallel | 4.60 (1.73, 7.95) | 0.47 (−0.07, 0.88) | 7.62 | 0 | 2 | 31 |
| Fall model | Delpierre | 8.05 (3.54, 13.65) | 0.33 (0.06, 0.64) | N/A | N/A | N/A | N/A |

^aIn columns 3 and 4 the figures in parentheses are the 5 and 95 percentiles of the value from all the weather stations. For 128 stations with more than 5 years' meteorological data, columns 6, 7 and 8 show the numbers of stations for which the model is considered best in comparison to the other two; the number of station where the difference between the AIC_c of the model and the best model is less than 2 ($\Delta AIC_c < 2$) and less than 6 ($\Delta AIC_c < 6$).

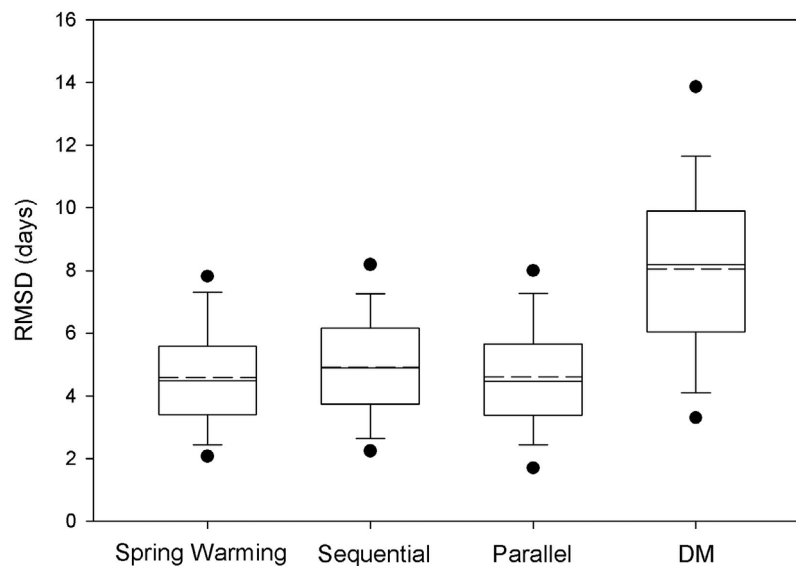


Figure 5. Box plot of the RMSD distribution of three budburst models (the SW, SEQ, and PAR models) and Delpierre model. The solid lines in the middle of the box are the median values. The dashed lines are mean values. The black dots are 5% and 95% percentile of the RMSDs of each model. The whiskers are the standard deviation of the RMSD from all the stations.

(Figure 5, Figure 6). However, the AIC_c scores of all the stations (Table 3, a total of 128 stations, 9 stations with only 5 years' data were excluded because it will cause a zero denominator) support the spring warming model, only for two stations can the PAR model be considered to be approximately equal to spring warming model ($\Delta AIC_c < 2$), and more than 3/4 (100/128 and 97/128 for SEQ and PAR, respectively) of the stations are less than 20 times to be the true model. Model predictability varies across the region (Figures 7a–7c). Stations with highest RMSD were mainly distributed in the coastal area and the low elevation stations near the metropolitan Boston. The senescence model (i.e., DM) showed a higher overall and averaged RMSD than the budburst models though similarly showed high RMSD stations were mainly in the coastal area. Stations with the lowest RMSD were distributed in the inland high elevation areas. The averaged ME for the four tested models were listed in Table 3. The spring warming model was most efficient comparing to a null model which only represents the averaged start/end of season at each weather station. Both the sequential and parallel models were on average better than a null model but in some specific weather stations the ME is below zero. In a similar way to RMSD, the SW and PAR models have better performance than the SEQ model. DM had a better performance than the null model both in average and for each weather station.

[23] The model parameter distributions are shown in Figure 8. Most of the stations have a base heating temperature (T_{heat}) requirement of 3°–5°C (Figures 8a, 8d, 8h). T_{heat} for the spring warming model is generally skewed towards zero while those for SEQ and PAR are uniform. Spring warming models mostly start the heat accumulation (accumulation of GDD) at DOY 80–100 (about 20 March to 10 April) (Figure 8b). The base chilling temperature (T_{chill}) requirement for the SEQ and PAR models have a peak at ~3°C, which is more conservative than T_{heat} . The critical

forcing temperature for the three budburst models are mostly in the range of 200°–400°C (Figures 8c, 8f, 8j). Two 2-phase models have a base chilling temperature requirement of 0°–2°C (Figures 8g, 8k). For the senescence model, the threshold photoperiod is mostly between 11 and 13 hours, which for the study area occurs between September and mid-October (Figure 8l).

3.3. Retrospect Analysis Using 50 Years of Climate Data

[24] We found a statistically significant advancement of SOS in New England since 1960 (Table 4 and Figure 9) of an average of 0.143 days per year ($p = 0.015$). The advancement rate varies from station to station from 2.4 days/decade to 0.5 days per decade (Figure 10a). The stations with earlier SOS contribute to the lower envelop while those with later SOS contributes to the upper envelop in Figure 9a. On the contrary, EOS did not show a statistically significant delay or advance in the study area ($p = 0.3660$). This is basically a consequence of some stations showing an advance (~53%) while the others showing a delay (~47%) (Figure 10b). Combined together the trend in GSL is not statistically significant (although the slope is positive: 0.0638, $p = 0.4148$). Similar to EOS, the rate of change for GSL varies with location, with the majority of the stations (~70%) showing a lengthening of GSL (Figure 10c).

4. Discussion

4.1. Uncertainty of the Remote Sensing Observations

[25] The uncertainty analysis suggested that the remotely sensed phenology algorithms could possibly capture both the spring and the fall canopy change. The curve-fitting process starts with the screening of cloud-contaminated data points. In addition to the cloud tags in the MODIS reflectance products, we utilized the Savitzky-Golay filter, and the

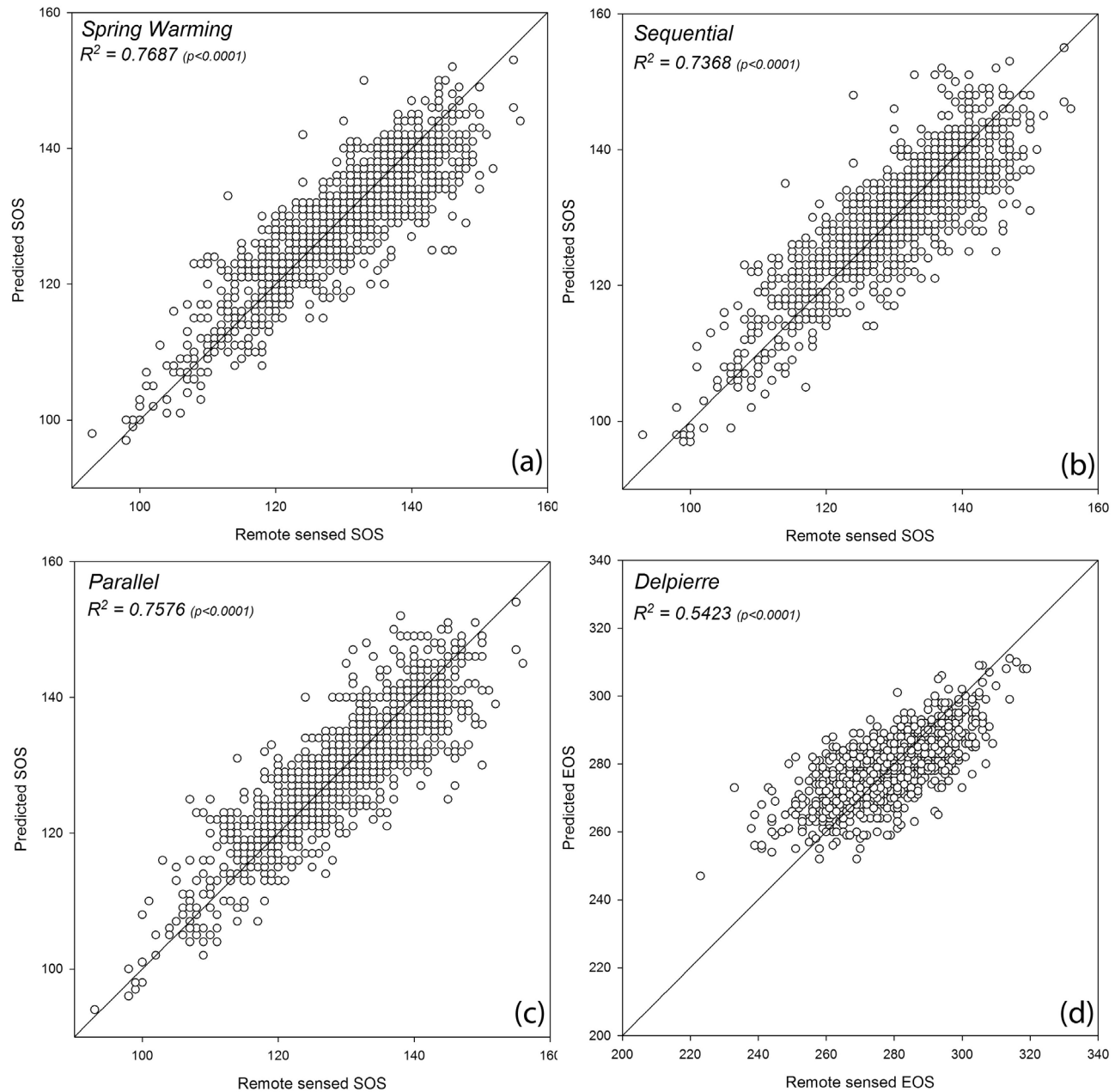


Figure 6. Scatter plots of the predicted phenophases (SOS/EOS) using models at each weather station and the remotely sensed SOS/EOS. Each plot shows the R^2 and p -value of the linear regression. The solid line is the 1:1 line; (a) spring warming model; (b) sequential model; (c) parallel model; and (d) Delpierre model.

double-logistic function to smooth the EVI2 time series. This algorithm effectively constrained the shape of the curve even when there were spikes in the winter (Figure 2b). Most of the EVI2 data points are within the 95% confidence interval of the fitted curve, especially for the points in the spring and fall.

[26] The remotely sensed SOS tracked the interannual variations of the ground-based phenological metrics (Figure 4a). The remotely sensed SOS is more correlated to the leaf expansion than budburst, probably because that the increase of leaf area is a stronger factor for the vegetation signal measured by the satellite sensor [Carlson and Ripley,

1997]. The relationships between fall phenological metrics and remotely sensed EOS are weaker. The uncertainties in the comparison are due to (1) the different scales of the observations (ground observation track vs. satellite pixels) and (2) the diverse phenological strategies of different species within the remotely sensed pixel [Steltzer and Post, 2009]. For (1), Digital camera-based phenological observations could potentially bridge the gap between satellite and ground observations [Hufkens et al., 2012]. For (2), we found that the variations of spring phenological metrics are smaller than those of fall phenological metrics (ground observations): no statistically significant difference was

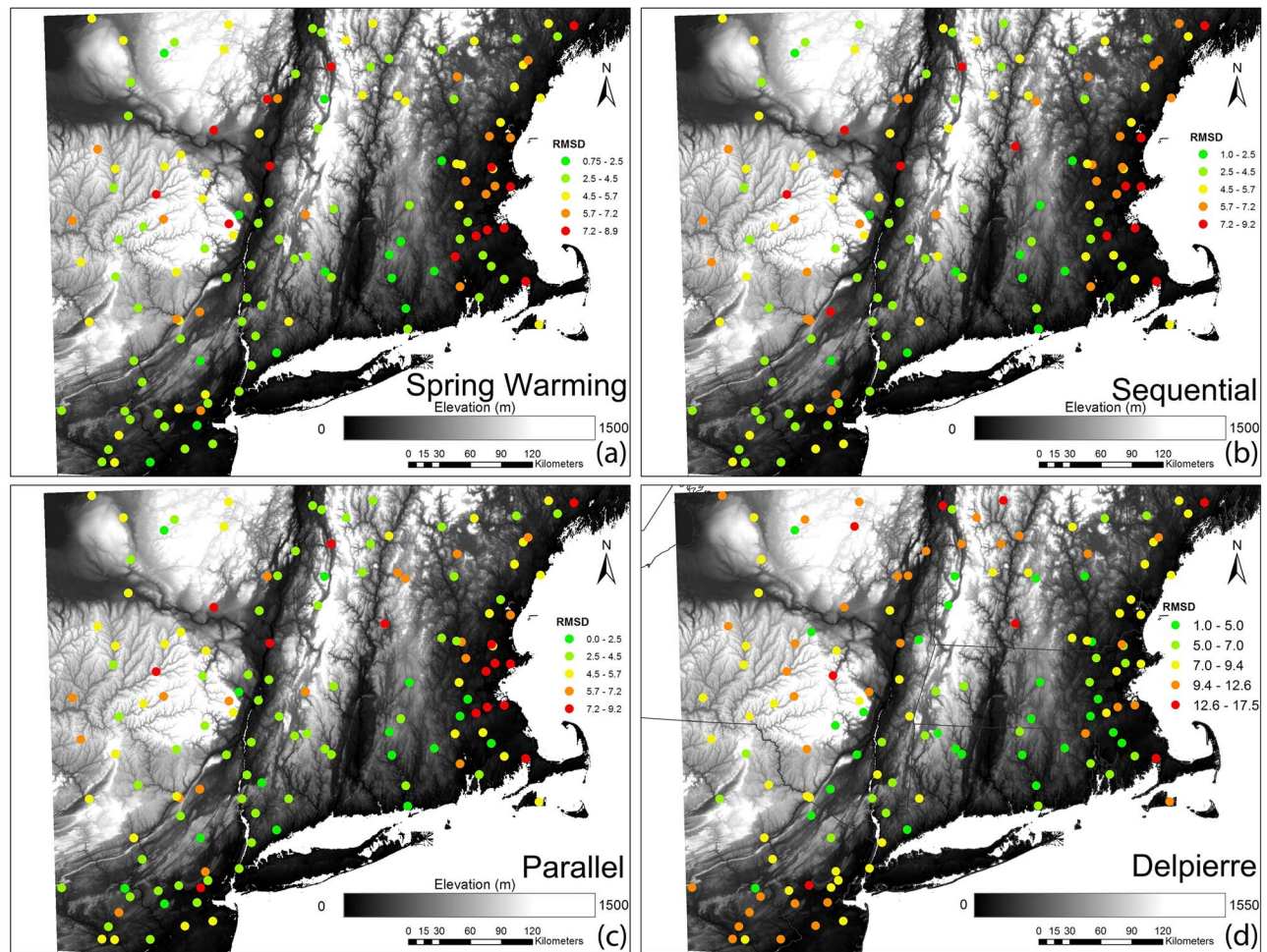


Figure 7. The RMSD of the four models at each individual weather station: (a) spring warming model; (b) sequential model; (c) parallel model; and (d) Delpierre model.

observed between species and individuals in the spring (data not shown); however, the differences between red oak and red maple in terms of LCOLOR and LFALL were statistically significant ($p = 0.0000$). Red maple changed leaf color ~ 2 weeks earlier than the red oaks (interannual average: DOY 274 vs. 290, t -test: $p = 0.0000$), and red oaks often retain their senescencing leaves much longer (O'Keefe, personal communication). This could increase the growing season length and delay the EOS calculated from remote sensing data, compared to the early senescencing species. However, the observed EOS is within the standard deviation of LCOLOR (Figure 4c).

4.2. Models Hypothesis and Comparison

[27] When applying species-level models to the regional level, one important question is whether the model parameters vary across the study area. Each species has its own phenology model parameters when tested against ground observations [Chuine, 2000; Delpierre, et al., 2009; Migliavacca et al., 2012; Richardson and O'Keefe, 2009]. Fisher et al. [2007] tested several hypotheses by applying the phenology model at the regional level. They refuted the hypotheses that forests in different locations share some common parameters (T_{heat} and t_0) while allowing other

parameter (F^*) to vary. The only remaining possible hypothesis is that the model parameters are station-specific and stratified by forest type. Modeling work based on ground observations supported this hypothesis: Richardson et al. [2006] found an overestimation of spring phenology in Harvard Forest when they used well-fitted spring phenology models at the more northerly Hubbard Brook Forest. Our study further supported the hypothesis by fitting the model at each individual station, and thus improved on the previous study of Fisher et al. [2007] for its short time span of good quality remote sensing data. Since the climate stations differ in the species type and composition, we observed that the phenology model parameters vary from station to station. This method could be extended to the areas without meteorological stations using only remotely sensed data (e.g., MODIS) and gridded climate data [Picard et al., 2005]. The phenology models we used are the models for a mixture of vegetation species, which may have different phenology strategies and thus model parameters. Although limited by the species mixture, we found that the parameters were within the range of the other studies and theories based on the controlled experiments [Chuine, 2000; Kramer, 1994]. An average base temperature for SOS in this region is $\sim 2.74^\circ \text{C}$, and a start date of 79, which is within the range

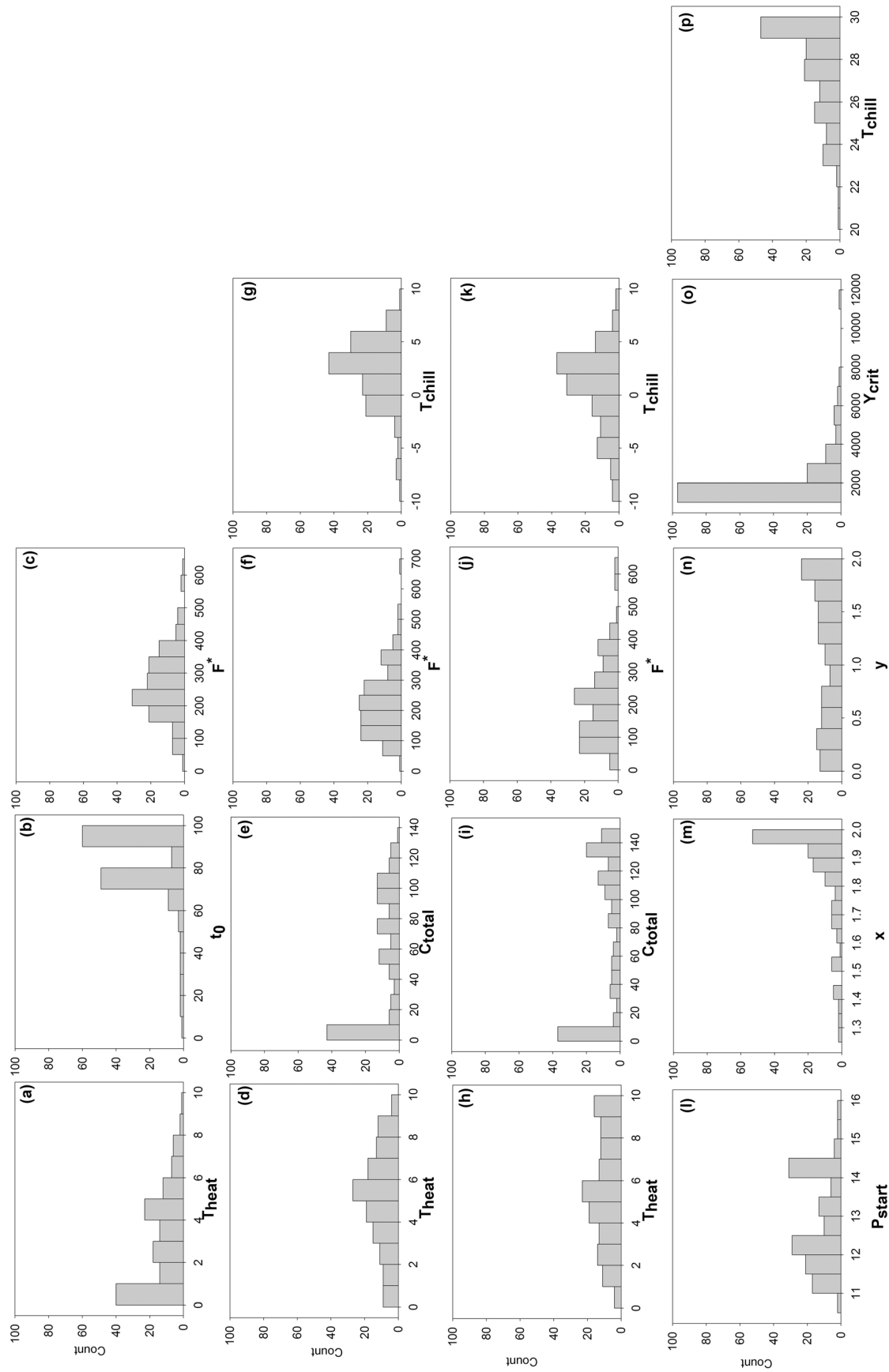


Figure 8. Histograms of parameter value of the four models: (a–c) SW; (d–g) SEQ; (h–k) PAR; (l–p) DM. Bin number = 10. Refer to Table 1 for the acronyms.

Table 4. Slope of the Linear Regression Using the Average SOS, EOS and GSL of All Stations With ME > 0.4 From 1960 to 2010 and the p -Value of the Slope of the Average Dates^a

| Phenophases | Slope | p -Value |
|-------------|-------------------------|------------|
| SOS | -0.143 (-0.243, -0.052) | 0.0152 |
| EOS | -0.078 (-0.488, 0.142) | 0.3660 |
| GSL | 0.065 (-0.341, 0.353) | 0.4148 |

^aIn column 2 the figures in parentheses are the 5th and 95th percentiles of the value.

of others' work [Chuine *et al.*, 2000; Hänninen, 1990]. Future work needs to address the effect of species composition on both the remotely sensed phenology (especially SOS and EOS) and phenology models at regional and global scales. Data fusion using Landsat TM (resolution: 30m) and MODIS (resolution: 500m) [Fisher and Mustard, 2007; Zhu *et al.*, 2010] could be used to track the vegetation dynamics at the scale comparable to the ground-level Forest Inventory Analysis forest plots, which provide species-composition information that can be used to understand the spatial variation of phenology model parameters.

[28] Model evaluation based on performance measured with RMSD, AIC_c and ME suggested that among the selected budburst models, the SW model was the best. We suggested that model evaluation should not only be based on the goodness-of-fit, but also the model complexity (i.e., the number of parameters). When only consider RMSD, the three budburst models are close. The RMSD of the SEQ model is statistically significantly higher than those of the SW and the PAR models ($p = 0.0000$) while the SW and the PAR are not statistically significantly different ($p = 0.9069$). However, when AIC_c and ME are considered in the evaluation, we found that the SW is the best choice. AIC_c values from all stations support the SW model (Table 3). Only fewer than 1% of the stations support that the PAR model is close to the best model – the SW. Comparing to the SW, both the SEQ and the PAR models have more parameters, which in AIC_c will be penalized for their more complex model structure. In addition, the averaged ME is higher for the SW model than that for the PAR model. In some weather stations, the ME of the PAR model is lower than 0 (Table 3, meaning less effective than a null model), suggesting that the PAR model only works for limited areas. This might be due to the structure of the budburst models: the SW model implicitly assumes that chilling requirements in the winter are always fulfilled, while the SEQ and the PAR models need to fulfill a certain form of chilling requirement otherwise the budburst might be delayed. Previous work based on satellite and climate data found that in North America, from 40°N northward the chilling requirements are always fulfilled [Zhang *et al.*, 2007]. Thus, additional parameters in the model structure (i.e., parameters for the chilling part) are not necessary. In addition, a comparison of the SW, SEQ, and PAR models using ground observation data in Harvard Forest also suggests that under current climate (for period 1990–2006), the SW model is still the best choice for 1/3 of the species, and the PAR model is the second choice [Richardson and O'Keefe, 2009]. However, the SEQ and the PAR models might become better in the future as the winter temperature in Northeast US is projected to increase about 2.9°C by 2100 comparing to 1961–1990 even under lower

emission scenarios [Hayhoe *et al.*, 2007], which might cause an unmet of winter chilling requirement. For the senescence model, the DM model showed an RMSD of ~1 week, which is higher than that of budburst models, suggesting that

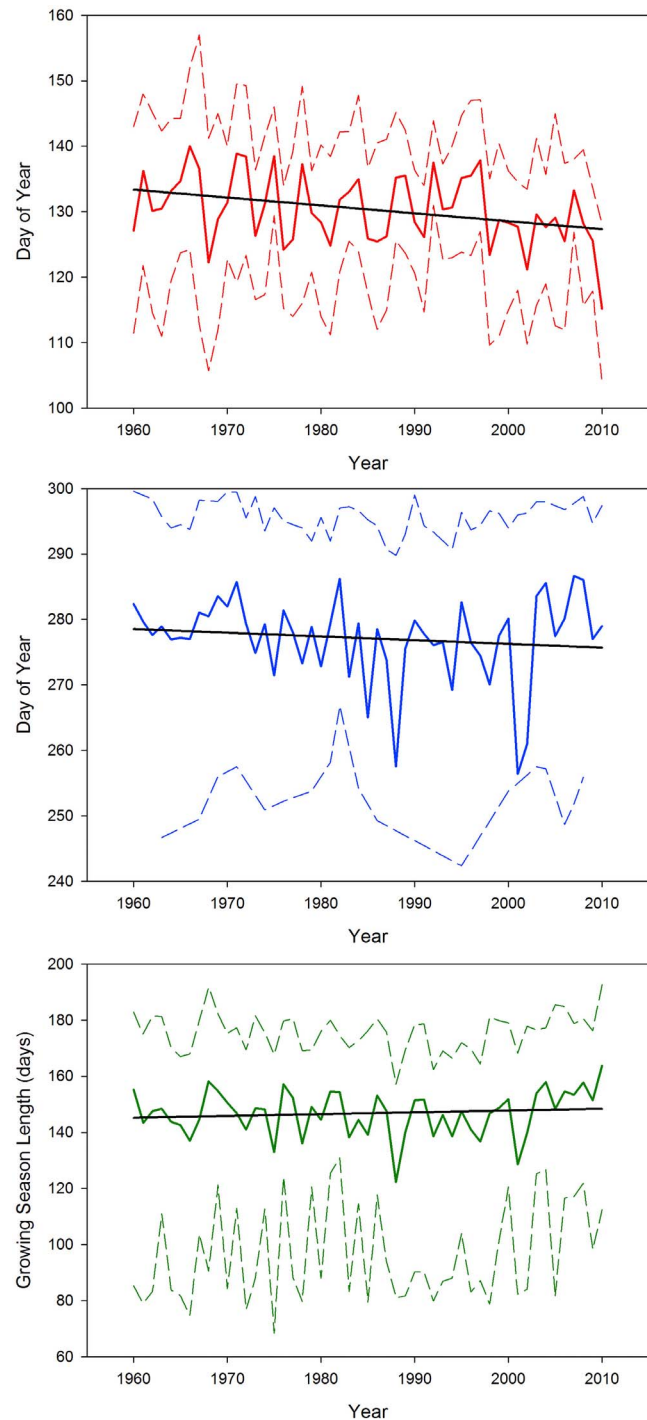


Figure 9. Retrospective analysis using the weather station data from 1960 to 2010 and the calibrated spring warming model (for SOS) and Delpierre model (for EOS). The solid color lines are the averaged phenophases from all the weather stations while the solid black lines are the linear regression (see Table 4). The dashed lines are the 5% and 95% percentile of the phenophases from all the selected stations.

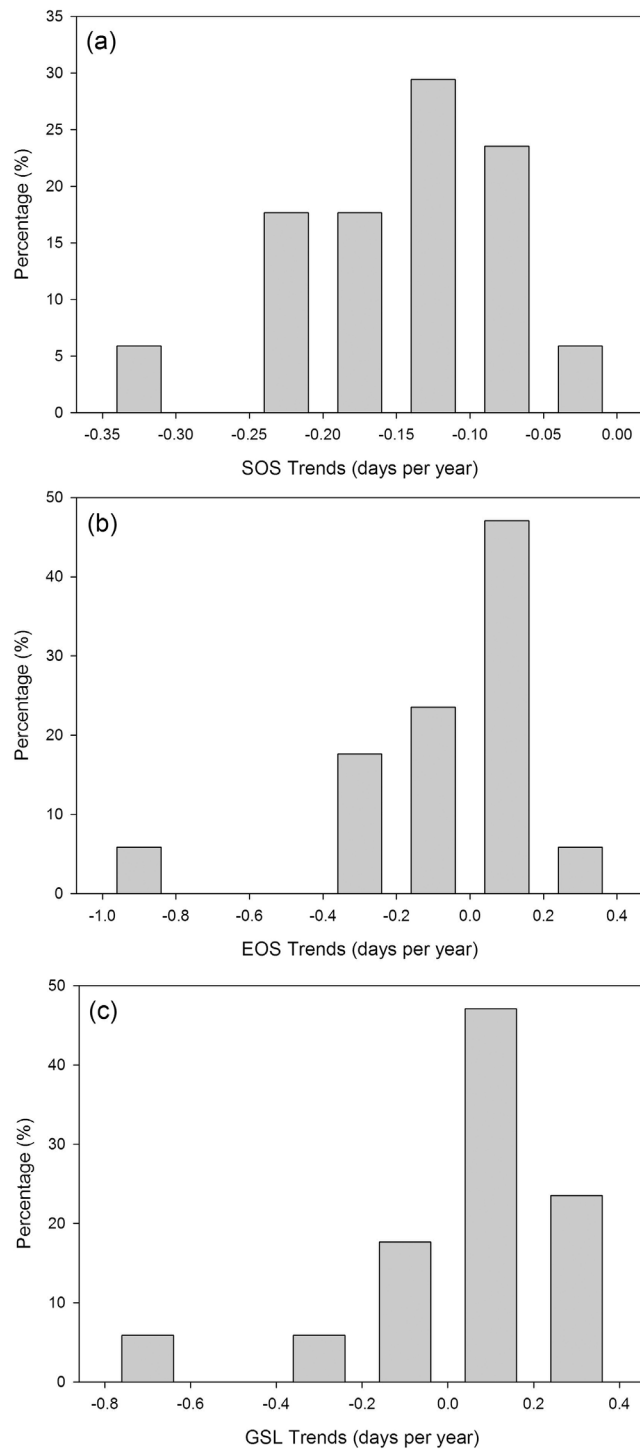


Figure 10. Histograms of the trends of phenophases from 1960 to 2010 in retrospective analysis. (a) start of season; (b) end of season; and (c) growing season length.

additional factors might contribute to the variance [Vitasse *et al.*, 2011].

4.3. Environmental Factors in the Phenology Models

[29] We only used temperature as the driver in our budburst models, since it is considered to be the dominant driver in spring phenology. However, the SW model implicitly

includes the photoperiod requirement by allowing the start date of heat accumulation (t_0) to change [Migliavacca *et al.*, 2012]; the start date in this region are mainly between 20 March and 10 April (Figure 8b), during which the day length is about 12–13 hours in Harvard Forest. In the spring, although some early succession species such as beech are opportunists that will respond to year-to-year variation of temperature, late succession species such as oak are adapted to the local change and are more responsive to the invariant environmental factors like photoperiod [Lechowicz, 1984; Polgar and Primack, 2011]. Recently, a budburst model explicitly includes both temperature and photoperiod as the drivers, and showed a lower RMSD than the traditional the SW models when tested against ground-observed apple blossom data [Blümel and Chmielewski, 2012]. Photoperiod could be considered as a potential parameter in budburst model in the future, although it needs to be examined if the decrease in RMSD is the result of the inclusion of photoperiod mechanism or the additional parameters (In which case, AIC_c should be used).

[30] For senescence, temperature, on average, is suggested to be more important than photoperiod in controlling the senescence process [Vitasse *et al.*, 2011]. We found a similar result that the temperature parameter x is significantly higher than the photoperiod parameter y (t -test, $p = 0.0000$) (Figure 11). However, we found that the relative importance of temperature and photoperiod varies across the region, and shows no clear relationship with the latitude or elevation (Figure 12). This is possibly due to the species composition: Delpierre *et al.* [2009] found that the senescence of *Quercus* is not modulated by photoperiod (parameter $y = 0$) while the senescence of *Fagus* is controlled by both temperature and photoperiod (parameter $x = 2$, $y = 2$).

[31] The spatial distributions of RMSD of the four models showed a coastal high-inland low trend. The high RMSD in the coastal region might be caused by the satellite sensor drift, ocean proximity and soil type [Fisher *et al.*, 2007; Motzkin *et al.*, 2002]. For budburst models, RMSDs are highest near urban areas such as Boston. This might be partly due to the urban/vegetation mixture leading to noise in the seasonal trajectory of vegetation signal [Fisher and Mustard, 2007]. In addition, anthropogenic effect such as N deposition and water availability change (not parameterized in the models) [Sherry *et al.*, 2007] might result in the diverse response of plant phenology [Cleland *et al.*, 2006], leading to less accurate models.

4.4. Phenological Trends in New England From 1960 to 2010

[32] The trends of the SOS, EOS, and GSL are divergent in direction, amplitude and the significance. The averaged advance rate of SOS in New England of 1.4 days per decade (from 1960 to 2010, with a 5% and 95% percentile of 0.5 days per decade and 2.4 days per decade) found in the retrospective analysis is close to the findings of other analysis from the U.S: field observations and retrospective analysis using the Hubbard Brook Experimental Forest (within the New England) data suggest the rate is 1.8 days per decade from 1957 to 2004 [Richardson *et al.*, 2006]. Results from a terrestrial biosphere model (ORCHIDEE) found that a 1.6 days per decade (1980–2002) increase in northern hemisphere start of season [Piao *et al.*, 2007]. The lilac

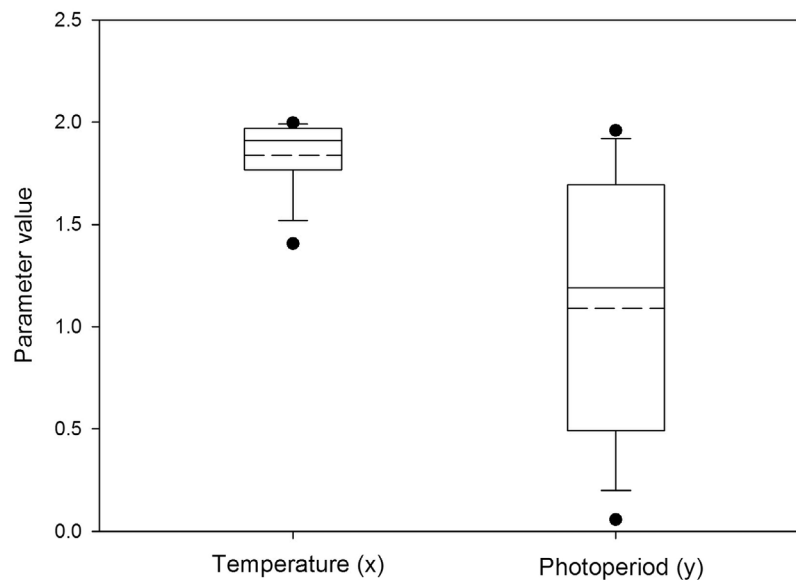


Figure 11. Box plot of the distribution of values of parameters x and y in DM. These parameters are the indicators of relative importance of temperature (x) and photoperiod (y). The solid lines in the middle of the box represent the median value. The dashed lines indicate the mean value. The two whiskers are the standard deviation. The black dots are 5% and 95% percentile. The p value for the t -test ($p = 0.000000$) suggested temperature has significantly different effects on the senescence stage than photoperiod.

records extending from 1956 to 2003 were incorporated in a temperature-driven spring index, which suggests an advancement of 1.2 days/per decade [Schwartz *et al.*, 2006]. Remote sensing data from NOAA/AVHRR suggested that a 7.7 days advancement in the U.S. temperate and boreal forest in 18 years (1981–1999), i.e., 4.27 days per decade [Zhou *et al.*, 2001]. Notice that this analysis was using a coarser resolution satellite data (8km) and the result is an average across the entire latitudinal strip. Using the same dataset (NOAA/AVHRR) but a longer time span (1982–2008), Zhu, *W.*, *et al.* [2012] found that delayed EOS rather than advanced SOS dominated the vegetation phenological shift in North America (35°N–70°N). These discrepancies could be a result of the temporal scale (in our paper the retrospective analysis used dataset from 1960 to 2010) and spatial scale (New England (40°N–44°N) vs. the entire North America). Even between 1982 and 2008, phenological shifts for SOS and EOS might not be invariant: the amplitudes of SOS advance and EOS delay were larger in 1982–1999 than those in 2000–2008 [Jeong *et al.*, 2011]. Our results suggest that by combining remote sensing and meteorological data, instead of a single site, we could potentially reconstruct the phenology of deciduous forests in the past several decades.

[33] Due to the lack of long-term ground observation data of senescence in North America, we were not able to compare with the other results in the same area. However, in the similar latitude in Europe, Menzel *et al.* [2006] found a diverse fall season response to temperature variations, and only 3% of the species investigated were significantly delayed in autumn, including *Fagus* (+0.6 days/decade) and *Quercus* (+1.0 days/decade) during 1951–1999 [Delpierre *et al.*, 2009]. To improve the ability of phenology model, the characterization of vegetation senescence is an important next step [Richardson *et al.*, 2012]. Overall, the agreement

with different scales of phenology data suggests the feasibility of the models being applied to the regional scale.

5. Conclusion

[34] Changes in the vegetation phenology may be an indicator of climate change. Species-level phenology models are considered to be more efficient than the phenology models used in the terrestrial biosphere models when tested against ground observation data [Migliavacca *et al.*, 2012]. Yet, species-level phenology models are rarely examined in a regional context, where remote sensing provides phenological observations covering large areas. Our results suggest among the three budburst model, the simplest model—the spring warming model—is the best: the model evaluation using AIC_c, RMSD and ME support the SW model instead of the SEQ and PAR models (and a null model). Similarly, the DM model was better than the fall null model at predicting the occurrence of senescence. The DM model parameters also suggested that temperature is the main driver of senescence, and photoperiod is of the second importance. We also found a statistically significant advancement of the SOS in New England (averaged advancement is 0.143 days per year) using the spring warming model and the magnitude of advancement varies from station to station. However, no significant advance or delay was observed for the EOS and the GSL in this region over the period of 1960 to 2010. Our findings suggest that species-level phenology models can be parameterized using satellite and meteorological data to construct vegetation phenology at regional scale, which can be extended to areas without meteorological stations where only remote sensing data (e.g., MODIS) and gridded climate data are available. This offers a method to improve the phenology models and support their incorporation into the terrestrial biosphere models. In addition, these results

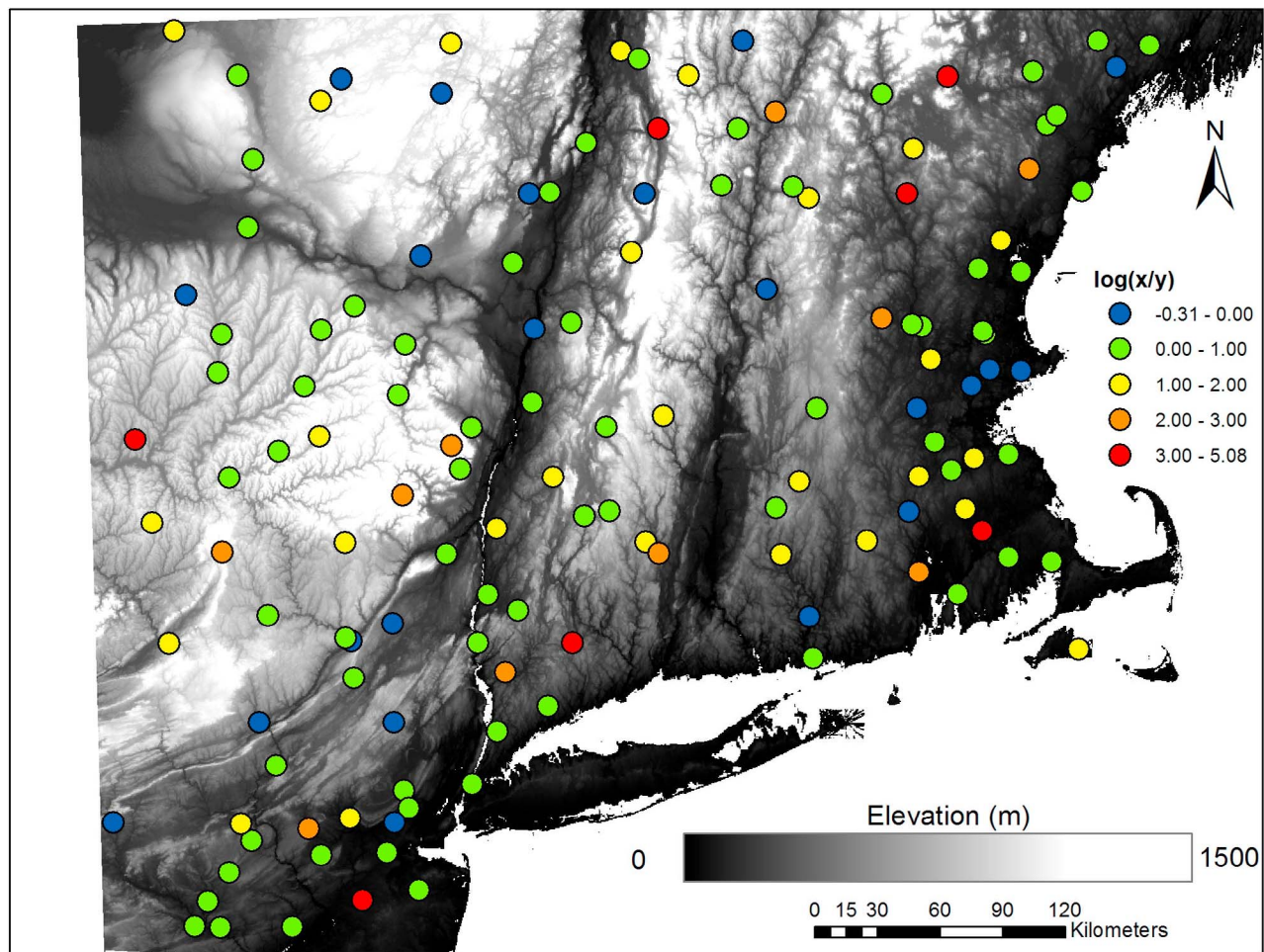


Figure 12. The relative strength of temperature (x) and photoperiod (y), calculated as $\log(x/y)$. A negative $\log(x/y)$ suggests $x < y$; a positive value suggests $x \geq y$.

suggest the possibility that the species-level models at the regional level can be used to track plants' response to past climate and predict the response to the future climate. Future research needs to address the effect of species composition on remotely sensed phenology. Digital-camera-based phenology observations could play an important role [Richardson, 2008; Richardson *et al.*, 2009; see also Hufkens *et al.*, 2012] in understanding how the diverse phenological strategies of different species could affect the remotely sensed phenology. Forest Inventory Analysis (FIA) dataset [Zhu, K., *et al.*, 2012] might help to give the detailed information of species composition of part of the area. Recent efforts to use spectroscopic method and LIDAR in tropical forest for vegetation classification could help to establish a regional-scale species distribution map [Asner and Martin, 2008, 2009]. In addition, efforts to understand the driving factors of senescence would help to improve the senescence modeling.

[35] **Acknowledgments.** We thank Dennis Baldocchi, Nicolas Delapierre, Johanna Schmitt, the Associate Editor, and an anonymous reviewer for constructive comments on an earlier version of this manuscript. We thank John O'Keefe for providing the Harvard Forest phenological data and for commenting on our revised manuscript. This research was supported by the Brown University–Marine Biological Laboratory graduate

program in Biological and Environmental Sciences, Brown–ECI phenology working group, and Brown Office of International Affairs Seed Grant on phenology.

References

- Akaike, H. (1974), A new look at the statistical model identification, *IEEE Trans. Autom. Control*, 19(6), 716–723, doi:10.1109/TAC.1974.1100705.
- Asner, G. P., and R. E. Martin (2008), Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels, *Remote Sens. Environ.*, 112(10), 3958–3970, doi:10.1016/j.rse.2008.07.003.
- Asner, G. P., and R. E. Martin (2009), Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests, *Front. Ecol. Environ.*, 7(5), 269–276, doi:10.1890/070152.
- Blümel, K., and F.-M. Chmielewski (2012), Shortcomings of classical phenological forcing models and a way to overcome them, *Agric. For. Meteorol.*, 164, 10–19, doi:10.1016/j.agrformet.2012.05.001.
- Bradshaw, W. E., and C. M. Holzapfel (2006), Evolutionary response to rapid climate change, *Science*, 312(5779), 1477–1478, doi:10.1126/science.1127000.
- Burnham, K. P., and D. R. Anderson (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed., Springer, Berlin.
- Cannell, M. G. R., and R. I. Smith (1983), Thermal time, chill days and prediction of budburst in *Picea sitchensis*, *J. Appl. Ecol.*, 20(3), 951–963, doi:10.2307/2403139.
- Carlson, T. N., and D. A. Ripley (1997), On the relation between NDVI, fractional vegetation cover, and leaf area index, *Remote Sens. Environ.*, 62(3), 241–252, doi:10.1016/S0034-4257(97)00104-1.

- Chen, J., P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh (2004), A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter, *Remote Sens. Environ.*, 91(3-4), 332–344, doi:10.1016/j.rse.2004.03.014.
- Chuine, I. (2000), A unified model for budburst of trees, *J. Theor. Biol.*, 207, 337–347, doi:10.1006/jtbi.2000.2178.
- Chuine, I., P. Cour, and D. D. Rousseau (1998), Fitting models predicting dates of flowering of temperate-zone trees using simulated annealing, *Plant Cell Environ.*, 21(5), 455–466, doi:10.1046/j.1365-3040.1998.00299.x.
- Chuine, I., P. Cour, and D. D. Rousseau (1999), Selecting models to predict the timing of flowering of temperate trees: Implications for tree phenology modelling, *Plant Cell Environ.*, 22(1), 1–13, doi:10.1046/j.1365-3040.1999.00395.x.
- Chuine, I., G. Cambon, and P. Comtois (2000), Scaling phenology from the local to the regional level: Advances from species-specific phenological models, *Global Change Biol.*, 6, 943–952, doi:10.1046/j.1365-2486.2000.00368.x.
- Churkina, G., D. Schimel, B. H. Braswell, and X. Xiao (2005), Spatial analysis of growing season length control over net ecosystem exchange, *Global Change Biol.*, 11(10), 1777–1787, doi:10.1111/j.1365-2486.2005.00102.x.
- Cleland, E. E., N. R. Chiariello, S. R. Loarie, H. A. Mooney, and C. B. Field (2006), Diverse responses of phenology to global changes in a grassland ecosystem, *Proc. Natl. Acad. Sci. U. S. A.*, 103(37), 13,740–13,744, doi:10.1073/pnas.0600815103.
- Cleland, E., I. Chuine, A. Menzel, H. Mooney, and M. Schwartz (2007), Shifting plant phenology in response to global change, *Trends Ecol. Evol.*, 22(7), 357–365, doi:10.1016/j.tree.2007.04.003.
- Cogbill, C. V., J. Burk, and G. Motzkin (2002), The forests of presettlement New England, USA: Spatial and compositional patterns based on town proprietor surveys, *J. Biogeogr.*, 29(10-11), 1279–1304, doi:10.1046/j.1365-2699.2002.00757.x.
- Delpierre, N., E. Dufrêne, K. Soudani, E. Ulrich, S. Cecchini, J. Boé, and C. François (2009), Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France, *Agric. For. Meteorol.*, 149(6-7), 938–948, doi:10.1016/j.agrformet.2008.11.014.
- Dodson, R., and D. Marks (1997), Daily air temperature interpolated at high spatial resolution over a large mountainous region, *Clim. Res.*, 8(1), 1–20, doi:10.3354/cr008001.
- Dragoni, D., H. P. Schmid, C. A. Wayson, H. Potter, C. S. B. Grimmer, and J. C. Randolph (2011), Evidence of increased net ecosystem productivity associated with a longer vegetated season in a deciduous forest in south-central Indiana, USA, *Global Change Biol.*, 17(2), 886–897, doi:10.1111/j.1365-2486.2010.02281.x.
- Estrella, N., and A. Menzel (2006), Responses of leaf colouring in four deciduous tree species to climate and weather in Germany, *Clim. Res.*, 32(3), 253–267, doi:10.3354/cr032253.
- Fisher, J., and J. Mustard (2007), Cross-scalar satellite phenology from ground, Landsat, and MODIS data, *Remote Sens. Environ.*, 109, 261–273, doi:10.1016/j.rse.2007.01.004.
- Fisher, J., J. Mustard, and M. Vadboucoeur (2006), Green leaf phenology at Landsat resolution: Scaling from the field to the satellite, *Remote Sens. Environ.*, 100, 265–279, doi:10.1016/j.rse.2005.10.022.
- Fisher, J., A. Richardson, and J. Mustard (2007), Phenology model from surface meteorology does not capture satellite-based green-up estimations, *Global Change Biol.*, 13, 707–721, doi:10.1111/j.1365-2486.2006.01311.x.
- Fitter, A. H., and S. R. Fitter (2002), Rapid changes in flowering time in British plants, *Science*, 296, 1689–1691, doi:10.1126/science.1071617.
- Foster, D., B. Hall, S. Barry, S. Clayden, and T. Parshall (2002), Cultural, environmental and historical controls of vegetation patterns and the modern conservation setting on the island of Martha's Vineyard, USA, *J. Biogeogr.*, 29, 1381–1400, doi:10.1046/j.1365-2699.2002.00761.x.
- Häkkinen, R., T. Linkosalo, and P. Hari (1998), Effects of dormancy and environmental factors on timing of bud burst in *Betula pendula*, *Tree Physiol.*, 18(10), 707–712, doi:10.1093/treephys/18.10.707.
- Hänninen, H. (1990), Modelling bud dormancy release in trees from cool and temperate regions, *Acta Forestalia Fennica*, 213, 1–47.
- Hayhoe, K., C. P. Wake, T. G. Huntington, L. Luo, M. D. Schwartz, J. Sheffield, E. Wood, B. Anderson, J. Bradbury, and A. DeGaetano (2007), Past and future changes in climate and hydrological indicators in the US Northeast, *Clim. Dyn.*, 28(4), 381–407, doi:10.1007/s00382-006-0187-8.
- Hufkens, K., M. Friedl, O. Sonnentag, B. H. Braswell, T. Milliman, and A. D. Richardson (2012), Linking near-surface and satellite remote sensing measurements of deciduous broadleaf forest phenology, *Remote Sens. Environ.*, 117(0), 307–321, doi:10.1016/j.rse.2011.10.006.
- Hunter, A. F., and M. J. Lechowicz (1992), Predicting the timing of budburst in temperate trees, *J. Appl. Ecol.*, 29(3), 597–604, doi:10.2307/2404467.
- Jeong, S.-J., C.-H. Ho, H.-J. Gim, and M. E. Brown (2011), Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008, *Global Change Biol.*, 17(7), 2385–2399, doi:10.1111/j.1365-2486.2011.02397.x.
- Jiang, Z., A. R. Huete, K. Didan, and T. Miura (2008), Development of a two-band enhanced vegetation index without a blue band, *Remote Sens. Environ.*, 112(10), 3833–3845, doi:10.1016/j.rse.2008.06.006.
- Keskitalo, J., G. Bergquist, P. Gardeström, and S. Jansson (2005), A cellular timetable of autumn senescence, *Plant Physiol.*, 139(4), 1635–1648, doi:10.1104/pp.105.066845.
- Kramer, K. (1994), Selecting a model to predict the onset of growth of *Fagus sylvatica*, *J. Appl. Ecol.*, 31(1), 172–181, doi:10.2307/2404609.
- Landsberg, J. (1974), Apple fruit bud development and growth: Analysis and an empirical model, *Ann. Bot. (Lond.)*, 38(5), 1013–1023.
- Lechowicz, M. J. (1984), Why do temperate deciduous trees leaf out at different times? Adaptation and ecology of forest communities, *Am. Nat.*, 124(6), 821–842, doi:10.1086/284319.
- Markwardt, C. B. (2008), Non-linear least squares fitting in IDL with MPFIT, in *Proceedings: Astronomical Data Analysis Software and Systems XVIII, Quebec, Canada, ASP Conf. Ser.*, vol. 411, edited by D. Bohlender, P. Dowler, and D. Durand, pp. 251–254, Astron. Soc. of the Pacific, San Francisco, Calif.
- Menzel, A., et al. (2006), European phenological response to climate change matches the warming pattern, *Global Change Biol.*, 12(10), 1969–1976, doi:10.1111/j.1365-2486.2006.01193.x.
- Migliavacca, M., O. Sonnentag, T. F. Keenan, A. Cescatti, J. O'Keefe, and A. D. Richardson (2012), On the uncertainty of phenological responses to climate change, and implications for a terrestrial biosphere model, *Biogeosciences*, 9(6), 2063–2083, doi:10.5194/bg-9-2063-2012.
- Monteith, J. L., and M. H. Unsworth (2008), *Principles of Environmental Physics*, 3rd ed., Academic, San Diego, Calif.
- Morissette, J., et al. (2009), Tracking the rhythm of the seasons in the face of global change: Phenological research in the 21st century, *Front. Ecol.*, 7(5), 253–260, doi:10.1890/070217.
- Motzkin, G., S. C. Ciccarello, and D. R. Foster (2002), Frost pockets on a level sand plain: Does variation in microclimate help maintain persistent vegetation patterns?, *J. Torrey Bot. Soc.*, 129(2), 154–163, doi:10.2307/3088728.
- Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I – A discussion of principles, *J. Hydrol.*, 10(3), 282–290, doi:10.1016/0022-1694(70)90255-6.
- O'Keefe, J. (2000), Phenology of woody species, Harvard Forest Data Archive: HF003, available at <http://harvardforest.fas.harvard.edu:8080/exist/xquery/data.xq?id=hf003>, Harvard Univ., Cambridge, Mass.
- Partanen, J., V. Koski, and H. Hänninen (1998), Effects of photoperiod and temperature on the timing of bud burst in Norway spruce (*Picea abies*), *Tree Physiol.*, 18(12), 811–816, doi:10.1093/treephys/18.12.811.
- Pau, S., E. M. Wolkovich, B. I. Cook, T. J. Davies, N. J. B. Kraft, K. Bolmgren, J. L. Betancourt, and E. E. Cleland (2011), Predicting phenology by integrating ecology, evolution and climate science, *Global Change Biol.*, 17(12), 3633–3643, doi:10.1111/j.1365-2486.2011.02515.x.
- Peñuelas, J., and I. Fillela (2001), Responses to a warming world, *Science*, 294, 793–795, doi:10.1126/science.1066860.
- Peñuelas, J., T. Rutishauser, and I. Fillela (2009), Phenology feedbacks on climate change, *Science*, 324(5929), 887–888, doi:10.1126/science.1173004.
- Piao, S., P. Friedlingstein, P. Ciais, N. Viovy, and J. Demarty (2007), Growing season extension and its impact on terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades, *Global Biogeochem. Cycles*, 21, GB3018, doi:10.1029/2006GB002888.
- Picard, G., S. Quegan, N. Delbart, M. R. Lomas, T. Le Toan, and F. I. Woodward (2005), Bud-burst modelling in Siberia and its impact on quantifying the carbon budget, *Global Change Biol.*, 11(12), 2164–2176, doi:10.1111/j.1365-2486.2005.01055.x.
- Polgar, C. A., and R. B. Primack (2011), Leaf-out phenology of temperate woody plants: From trees to ecosystems, *New Phytol.*, 191(4), 926–941, doi:10.1111/j.1469-8137.2011.03803.x.
- Randerson, J. T., et al. (2009), Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, *Global Change Biol.*, 15(10), 2462–2484, doi:10.1111/j.1365-2486.2009.01912.x.
- Reich, P. B. (1995), Phenology of tropical forests: Patterns, causes, and consequences, *Can. J. Bot.*, 73(2), 164–174, doi:10.1139/b95-020.
- Richardson, A. D. (2008), Harvard Forest PhenoCam Images, Harvard Forest Data Archive:HF158, available at <http://phenocam.sr.unh.edu/webcam/>, Harvard Univ., Cambridge, Mass.
- Richardson, A. D., and J. O'Keefe (2009), Phenological differences between understory and overstory: A case study using the long-term Harvard Forest Records, in *Phenology of Ecosystem Processes: Applications in*

- Global Change Research*, edited by A. Noormets, pp. 87–117, Springer, New York, doi:10.1007/978-1-4419-0026-5_4.
- Richardson, A. D., A. Bailey, E. Denny, C. Martin, and J. O’Keefe (2006), Phenology of a northern hardwood forest canopy, *Global Change Biol.*, *12*, 1174–1188, doi:10.1111/j.1365-2486.2006.01164.x.
- Richardson, A. D., B. Braswell, D. Hollinger, J. Jenkins, and S. Ollinger (2009), Near-surface remote sensing of spatial and temporal variation in canopy phenology, *Ecol. Appl.*, *19*(6), 1417–1428, doi:10.1890/08-2022.1.
- Richardson, A. D., et al. (2010), Influence of spring and autumn phenological transitions on forest ecosystem productivity, *Philos. Trans. R. Soc. B*, *365*(1555), 3227–3246, doi:10.1098/rstb.2010.0102.
- Richardson, A. D., et al. (2012), Terrestrial biosphere models need better representation of vegetation phenology: Results from the North American Carbon Program Site Synthesis, *Global Change Biol.*, *18*(2), 566–584, doi:10.1111/j.1365-2486.2011.02562.x.
- Sarvas, R. (1974), Investigations on the annual cycle of development of forest trees: II. Autumn dormancy and winter dormancy, *Commun. Inst. For. Fenn.*, *84*, 1–101.
- Schaber, J., and F.-W. Badeck (2003), Physiology-based phenology models for forest tree species in Germany, *Int. J. Biometeorol.*, *47*(4), 193–201, doi:10.1007/s00484-003-0171-5.
- Schwartz, M. D. (1996), Examining the spring discontinuity in daily temperature ranges, *J. Clim.*, *9*(4), 803–808.
- Schwartz, M. D., R. Ahas, and A. Aasa (2006), Onset of spring starting earlier across the Northern Hemisphere, *Global Change Biol.*, *12*(2), 343–351, doi:10.1111/j.1365-2486.2005.01097.x.
- Sherry, R. A., X. Zhou, S. Gu, J. A. Arnone, D. S. Schimel, P. S. Verburg, L. L. Wallace, and Y. Luo (2007), Divergence of reproductive phenology under climate warming, *Proc. Natl. Acad. Sci. U. S. A.*, *104*(1), 198–202, doi:10.1073/pnas.0605642104.
- Steltzer, H., and E. Post (2009), Seasons and life cycles, *Science*, *324*, 886–887, doi:10.1126/science.1171542.
- Taylor, J. R. (1997), *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 2nd ed., Univ. Sci. Books, Sausalito, Calif.
- Vitasse, Y., S. Delzon, E. Dufrêne, J.-Y. Pontailler, J.-M. Louvet, A. Kremer, and R. Michalet (2009), Leaf phenology sensitivity to temperature in European trees: Do within-species populations exhibit similar responses?, *Agric. For. Meteorol.*, *149*(5), 735–744, doi:10.1016/j.agrformet.2008.10.019.
- Vitasse, Y., C. François, N. Delapierre, E. Dufrêne, A. Kremer, I. Chuine, and S. Delzon (2011), Assessing the effects of climate change on the phenology of European temperate trees, *Agric. For. Meteorol.*, *151*(7), 969–980, doi:10.1016/j.agrformet.2011.03.003.
- Zhang, X., M. Friedl, C. Schaaf, A. Strahler, J. Hodges, F. Gao, B. Reed, and A. Huete (2003), Monitoring vegetation phenology using MODIS, *Remote Sens. Environ.*, *84*, 471–475, doi:10.1016/S0034-4257(02)00135-9.
- Zhang, X., D. Tarpley, and J. Sullivan (2007), Diverse responses of vegetation phenology to a warming climate, *Geophys. Res. Lett.*, *34*, L19405, doi:10.1029/2007GL031447.
- Zhou, L., C. Tucker, R. Kaufmann, D. Slayback, N. Shabanov, and R. Myneni (2001), Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999, *J. Geophys. Res.*, *106*(D17), 20,069–20,083, doi:10.1029/2000JD000115.
- Zhu, K., C. W. Woodall, and J. S. Clark (2012), Failure to migrate: Lack of tree range expansion in response to climate change, *Global Change Biol.*, *18*(3), 1042–1052, doi:10.1111/j.1365-2486.2011.02571.x.
- Zhu, W., H. Tian, X. Xu, Y. Pan, G. Chen, and W. Lin (2012), Extension of the growing season due to delayed autumn over mid and high latitudes in North America during 1982–2006, *Global Ecol. Biogeogr.*, *21*(2), 260–271, doi:10.1111/j.1466-8238.2011.00675.x.
- Zhu, X., J. Chen, F. Gao, X. Chen, and J. G. Masek (2010), An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions, *Remote Sens. Environ.*, *114*(11), 2610–2623, doi:10.1016/j.rse.2010.05.032.